# 1 Growth Models with Exogenous Saving Rates (the Solow–Swan Model)

## 1.1 The Basic Structure

The first question we ask in this chapter is whether it is possible for an economy to enjoy positive growth rates forever by simply saving and investing in its capital stock. A look at the cross-country data from 1960 to 2000 shows that the average annual growth rate of real per capita GDP for 112 countries was 1.8 percent, and the average ratio of gross investment to GDP was 16 percent.[1] However, for 38 sub-Saharan African countries, the average growth rate was only 0.6 percent, and the average investment ratio was only 10 percent. At the other end, for nine East Asian "miracle" economies, the average growth rate was 4.9 percent, and the average investment ratio was 25 percent. These observations suggest that growth and investment rates are positively related. However, before we get too excited with this relationship, we might note that, for 23 OECD countries, the average growth rate was 2.7 percent—lower than that for the East Asian miracles—whereas the average investment ratio was 24 percent—about the same as that for East Asia. Thus, although investment propensities cannot be the whole story, it makes sense as a starting point to try to relate the growth rate of an economy to its willingness to save and invest. To this end, it will be useful to begin with a simple model in which the only possible source of per capita growth is the accumulation of physical capital.

Most of the growth models that we discuss in this book have the same basic general-equilibrium structure. First, households (or families) own the inputs and assets of the economy, including ownership rights in firms, and choose the fractions of their income to consume and save. Each household determines how many children to have, whether to join the labor force, and how much to work. Second, firms hire inputs, such as capital and labor, and use them to produce goods that they sell to households or other firms. Firms have access to a technology that allows them to transform inputs into output. Third, markets exist on which firms sell goods to households or other firms and on which households sell the inputs to firms. The quantities demanded and supplied determine the relative prices of the inputs and the produced goods.

Although this general structure applies to most growth models, it is convenient to start our analysis by using a simplified setup that excludes markets and firms. We can think of a composite unit—a household/producer like Robinson Crusoe—who owns the inputs and also manages the technology that transforms inputs into outputs. In the real world, production takes place using many different inputs to production. We summarize all of them into just three: physical capital $K(t)$, labor $L(t)$, and knowledge $T(t)$. The production

1. These data—from Penn World Tables version 6.1—are described in Summers and Heston (1991) and Heston, Summers, and Aten (2002). We discuss these data in chapter 12.

function takes the form

$$Y(t) = F[K(t), L(t), T(t)] \tag{1.1}$$

where $Y(t)$ is the flow of output produced at time $t$.

Capital, $K(t)$, represents the durable physical inputs, such as machines, buildings, pencils, and so on. These goods were produced sometime in the past by a production function of the form of equation (1.1). It is important to notice that these inputs cannot be used by multiple producers simultaneously. This last characteristic is known as *rivalry*—a good is *rival* if it cannot be used by several users at the same time.

The second input to the production function is labor, $L(t)$, and it represents the inputs associated with the human body. This input includes the number of workers and the amount of time they work, as well as their physical strength, skills, and health. Labor is also a *rival* input, because a worker cannot work on one activity without reducing the time available for other activities.

The third input is the level of knowledge or technology, $T(t)$. Workers and machines cannot produce anything without a *formula* or *blueprint* that shows them how to do it. This blueprint is what we call *knowledge or technology*. Technology can improve over time—for example, the same amount of capital and labor yields a larger quantity of output in 2000 than in 1900 because the technology employed in 2000 is superior. Technology can also differ across countries—for example, the same amount of capital and labor yields a larger quantity of output in Japan than in Zambia because the technology available in Japan is better. The important distinctive characteristic of knowledge is that it is a *nonrival good:* two or more producers can use the same formula at the same time.[2] Hence, two producers that each want to produce $Y$ units of output will each have to use a different set of machines and workers, but they can use the same formula. This property of nonrivalry turns out to have important implications for the interactions between technology and economic growth.[3]

2. The concepts of *nonrivalry* and *public good* are often confused in the literature. *Public goods* are *nonrival* (they can be used by many people simultaneously) and also *nonexcludable* (it is technologically or legally impossible to prevent people from using such goods). The key characteristic of knowledge is nonrivalry. Some formulas or blueprints are nonexcludable (for example, calculus formulas on which there are no property rights), whereas others are excludable (for example, the formulas used to produce pharmaceutical products while they are protected by patents). These properties of ideas were well understood by Thomas Jefferson, who said in a letter of August 13, 1813, to Isaac McPherson: "If nature has made any one thing less susceptible than all others of exclusive property, it is the actions of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of everyone, and the receiver cannot dispossess himself of it. Its peculiar character, too, is that no one possesses the less, because every other possesses the whole of it. He who receives an idea from me, receives instruction himself without lessening mine" (available on the Internet from the Thomas Jefferson Papers at the Library of Congress, lcweb2.loc.gov/ammem/mtjhtml/mtjhome.html).

3. Government policies, which depend on laws and institutions, would also affect the output of an economy. Since basic public institutions are nonrival, we can include these factors in $T(t)$ in the production function.

We assume a one-sector production technology in which output is a homogeneous good that can be consumed, $C(t)$, or invested, $I(t)$. Investment is used to create new units of physical capital, $K(t)$, or to replace old, depreciated capital. One way to think about the one-sector technology is to draw an analogy with farm animals, which can be eaten or used as inputs to produce more farm animals. The literature on economic growth has used more inventive examples—with such terms as *shmoos, putty,* or *ectoplasm*—to reflect the easy transmutation of capital goods into consumables, and vice versa.

In this chapter we imagine that the economy is closed: households cannot buy foreign goods or assets and cannot sell home goods or assets abroad. (Chapter 3 allows for an open economy.) We also start with the assumption that there are no government purchases of goods and services. (Chapter 4 deals with government purchases.) In a closed economy with no public spending, all output is devoted to consumption or gross investment,[4] so $Y(t) = C(t) + I(t)$. By subtracting $C(t)$ from both sides and realizing that output equals income, we get that, in this simple economy, the amount saved, $S(t) \equiv Y(t) - C(t)$, equals the amount invested, $I(t)$.

Let $s(\cdot)$ be the fraction of output that is saved—that is, the *saving rate*—so that $1 - s(\cdot)$ is the fraction of output that is consumed. Rational households choose the saving rate by comparing the costs and benefits of consuming today rather than tomorrow; this comparison involves preference parameters and variables that describe the state of the economy, such as the level of wealth and the interest rate. In chapter 2, where we model this decision explicitly, we find that $s(\cdot)$ is a complicated function of the state of the economy, a function for which there are typically no closed-form solutions. To facilitate the analysis in this initial chapter, we assume that $s(\cdot)$ is given exogenously. The simplest function, the one assumed by Solow (1956) and Swan (1956) in their classic articles, is a constant, $0 \le s(\cdot) = s \le 1$. We use this constant-saving-rate specification in this chapter because it brings out a large number of results in a clear way. Given that saving must equal investment, $S(t) = I(t)$, it follows that the *saving rate* equals the *investment rate*. In other words, the saving rate of a closed economy represents the fraction of GDP that an economy devotes to investment.

We assume that capital is a homogeneous good that depreciates at the constant rate $\delta > 0$; that is, at each point in time, a constant fraction of the capital stock wears out and, hence, can no longer be used for production. Before evaporating, however, all units of capital are assumed to be equally productive, regardless of when they were originally produced.

4. In an open economy with government spending, the condition is

$$Y(t) - r \cdot D(t) = C(t) + I(t) + G(t) + NX(t)$$

where $D(t)$ is international debt, $r$ is the international real interest rate, $G(t)$ is public spending, and $NX(t)$ is net exports. In this chapter we assume that there is no public spending, so that $G(t) = 0$, and that the economy is closed, so that $D(t) = NX(t) = 0$.

The net increase in the stock of physical capital at a point in time equals gross investment less depreciation:

$$\dot{K}(t) = I(t) - \delta K(t) = s \cdot F[K(t), L(t), T(t)] - \delta K(t) \tag{1.2}$$

where a dot over a variable, such as $\dot{K}(t)$, denotes differentiation with respect to time, $\dot{K}(t) \equiv \partial K(t)/\partial t$ (a convention that we use throughout the book) and $0 \leq s \leq 1$. Equation (1.2) determines the dynamics of $K$ for a given technology and labor.

The labor input, $L$, varies over time because of population growth, changes in participation rates, shifts in the amount of time worked by the typical worker, and improvements in the skills and quality of workers. In this chapter, we simplify by assuming that everybody works the same amount of time and that everyone has the same constant skill, which we normalize to one. Thus we identify the labor input with the total population. We analyze the accumulation of skills or human capital in chapter 5 and the choice between labor and leisure in chapter 9.

The growth of population reflects the behavior of fertility, mortality, and migration, which we study in chapter 9. In this chapter, we simplify by assuming that population grows at a constant, exogenous rate, $\dot{L}/L = n \geq 0$, without using any resources. If we normalize the number of people at time 0 to 1 and the work intensity per person also to 1, then the population and labor force at time $t$ are equal to

$$L(t) = e^{nt} \tag{1.3}$$

To highlight the role of capital accumulation, we start with the assumption that the level of technology, $T(t)$, is a constant. This assumption will be relaxed later.

If $L(t)$ is given from equation (1.3) and technological progress is absent, then equation (1.2) determines the time paths of capital, $K(t)$, and output, $Y(t)$. Once we know how capital or GDP changes over time, the growth rates of these variables are also determined. In the next sections, we show that this behavior depends crucially on the properties of the production function, $F(\cdot)$.

## 1.2 The Neoclassical Model of Solow and Swan

### 1.2.1 The Neoclassical Production Function

The process of economic growth depends on the shape of the production function. We initially consider the neoclassical production function. We say that a production function, $F(K, L, T)$, is *neoclassical* if the following properties are satisfied:[5]

5. We ignore time subscripts to simplify notation.

**1. Constant returns to scale.** The function $F(\cdot)$ exhibits constant returns to scale. That is, if we multiply capital and labor by the same positive constant, $\lambda$, we get $\lambda$ the amount of output:

$$F(\lambda K, \lambda L, T) = \lambda \cdot F(K, L, T) \quad \text{for all } \lambda > 0 \tag{1.4}$$

This property is also known as *homogeneity of degree one in K and L.* It is important to note that the definition of scale includes only the two rival inputs, capital and labor. In other words, we did not define constant returns to scale as $F(\lambda K, \lambda L, \lambda T) = \lambda \cdot F(K, L, T)$.

To get some intuition on why our assumption makes economic sense, we can use the following *replication argument.* Imagine that plant 1 produces $Y$ units of output using the production function $F$ and combining $K$ and $L$ units of capital and labor, respectively, and using formula $T$. It makes sense to assume that if we create an identical plant somewhere else (that is, if we *replicate* the plant), we should be able to produce the same amount of output. In order to replicate the plant, however, we need a new set of machines and workers, but we can use the same formula in both plants. The reason is that, while capital and labor are rival goods, the formula is a nonrival good and can be used in both plants at the same time. Hence, because technology is a nonrival input, our definition of returns to scale makes sense.

**2. Positive and diminishing returns to private inputs.** For all $K > 0$ and $L > 0$, $F(\cdot)$ exhibits positive and diminishing marginal products with respect to each input:

$$\begin{aligned} &\frac{\partial F}{\partial K} > 0, \qquad \frac{\partial^2 F}{\partial K^2} < 0 \\ &\frac{\partial F}{\partial L} > 0, \qquad \frac{\partial^2 F}{\partial L^2} < 0 \end{aligned} \tag{1.5}$$

Thus, the neoclassical technology assumes that, holding constant the levels of technology and labor, each additional unit of capital delivers positive additions to output, but these additions decrease as the number of machines rises. The same property is assumed for labor.

**3. Inada conditions.** The third defining characteristic of the neoclassical production function is that the marginal product of capital (or labor) approaches infinity as capital (or labor) goes to 0 and approaches 0 as capital (or labor) goes to infinity:

$$\begin{aligned} &\lim_{K\to 0}\left(\frac{\partial F}{\partial K}\right) = \lim_{L\to 0}\left(\frac{\partial F}{\partial L}\right) = \infty \\ &\lim_{K\to \infty}\left(\frac{\partial F}{\partial K}\right) = \lim_{L\to \infty}\left(\frac{\partial F}{\partial L}\right) = 0 \end{aligned} \tag{1.6}$$

These last properties are called *Inada conditions,* following Inada (1963).

**4. Essentiality.** Some economists add the assumption of *essentiality* to the definition of a neoclassical production function. An input is essential if a strictly positive amount is needed to produce a positive amount of output. We show in the appendix that the three neoclassical properties in equations (1.4)–(1.6) imply that each input is *essential* for production, that is, $F(0, L) = F(K, 0) = 0$. The three properties of the neoclassical production function also imply that output goes to infinity as either input goes to infinity, another property that is proven in the appendix.

**Per Capita Variables** When we say that a country is rich or poor, we tend to think in terms of output or consumption per person. In other words, we do not think that India is richer than the Netherlands, even though India produces a lot more GDP, because, once we divide by the number of citizens, the amount of income each person gets on average is a lot smaller in India than in the Netherlands. To capture this property, we construct the model in per capita terms and study primarily the dynamic behavior of the per capita quantities of GDP, consumption, and capital.

Since the definition of constant returns to scale applies to all values of $\lambda$, it also applies to $\lambda = 1/L$. Hence, output can be written as

$$Y = F(K, L, T) = L \cdot F(K/L, 1, T) = L \cdot f(k) \tag{1.7}$$

where $k \equiv K/L$ is capital per worker, $y \equiv Y/L$ is output per worker, and the function $f(k)$ is defined to equal $F(k, 1, T)$.[6] This result means that the production function can be expressed in *intensive form* (that is, in *per worker* or *per capita* form) as

$$y = f(k) \tag{1.8}$$

In other words, the production function exhibits no "scale effects": production per person is determined by the amount of physical capital each person has access to and, holding constant $k$, having more or fewer workers does not affect total output per person. Consequently, very large economies, such as China or India, can have less output or income per person than very small economies, such as Switzerland or the Netherlands.

We can differentiate this condition $Y = L \cdot f(k)$ with respect to $K$, for fixed $L$, and then with respect to $L$, for fixed $K$, to verify that the marginal products of the factor inputs are given by

$$\partial Y/\partial K = f'(k) \tag{1.9}$$

$$\partial Y/\partial L = f(k) - k \cdot f'(k) \tag{1.10}$$

6. Since $T$ is assumed to be constant, it is one of the parameters implicit in the definition of $f(k)$.

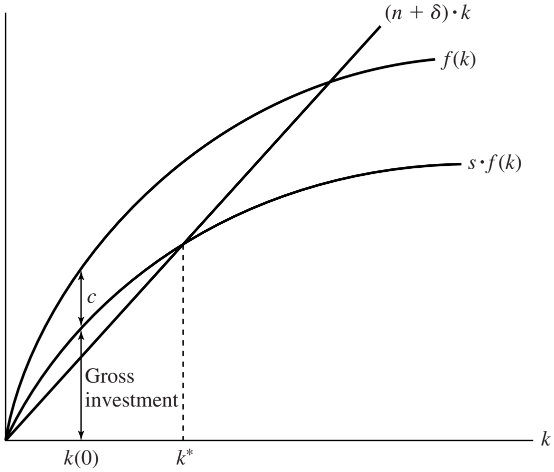


**Figure 1.1**
**The Solow–Swan model.** The curve for gross investment, $s \cdot f(k)$, is proportional to the production function, $f(k)$. Consumption per person equals the vertical distance between $f(k)$ and $s \cdot f(k)$. Effective depreciation (for $k$) is given by $(n+\delta) \cdot k$, a straight line from the origin. The change in $k$ is given by the vertical distance between $s \cdot f(k)$ and $(n+\delta) \cdot k$. The steady-state level of capital, $k^*$, is determined at the intersection of the $s \cdot f(k)$ curve with the $(n+\delta) \cdot k$ line.

The Inada conditions imply $\lim_{k\to 0}[f'(k)] = \infty$ and $\lim_{k\to\infty}[f'(k)] = 0$. Figure 1.1 shows the neoclassical production in per capita terms: it goes through zero; it is vertical at zero, upward sloping, and concave; and its slope asymptotes to zero as $k$ goes to infinity.

**A Cobb–Douglas Example** One simple production function that is often thought to provide a reasonable description of actual economies is the Cobb–Douglas function,[7]

$$Y = AK^{\alpha}L^{1-\alpha} \tag{1.11}$$

where $A > 0$ is the level of the technology and $\alpha$ is a constant with $0 < \alpha < 1$. The Cobb–Douglas function can be written in intensive form as

$$y = Ak^{\alpha} \tag{1.12}$$

7. Douglas is Paul H. Douglas, who was a labor economist at the University of Chicago and later a U.S. Senator from Illinois. Cobb is Charles W. Cobb, who was a mathematician at Amherst. Douglas (1972, pp. 46–47) says that he consulted with Cobb in 1927 on how to come up with a production function that fit his empirical equations for production, employment, and capital stock in U.S. manufacturing. Interestingly, Douglas says that the functional form was developed earlier by Philip Wicksteed, thus providing another example of Stigler's Law (whereby nothing is named after the person who invented it).

Note that $f'(k) = A\alpha k^{\alpha-1} > 0$, $f''(k) = -A\alpha(1-\alpha)k^{\alpha-2} < 0$, $\lim_{k\to\infty} f'(k) = 0$, and $\lim_{k\to 0} f'(k) = \infty$. Thus, the Cobb–Douglas form satisfies the properties of a neoclassical production function.

The key property of the Cobb–Douglas production function is the behavior of factor income shares. In a competitive economy, as discussed in section 1.2.3, capital and labor are each paid their marginal products; that is, the marginal product of capital equals the rental price $R$, and the marginal product of labor equals the wage rate $w$. Hence, each unit of capital is paid $R = f'(k) = \alpha Ak^{\alpha-1}$, and each unit of labor is paid $w = f(k) - k \cdot f'(k) = (1-\alpha) \cdot Ak^{\alpha}$. The capital share of income is then $Rk/f(k) = \alpha$, and the labor share is $w/f(k) = 1 - a$. Thus, in a competitive setting, the factor income shares are constant—independent of $k$—when the production function is Cobb–Douglas.

### 1.2.2 The Fundamental Equation of the Solow–Swan Model

We now analyze the dynamic behavior of the economy described by the neoclassical production function. The resulting growth model is called the Solow–Swan model, after the important contributions of Solow (1956) and Swan (1956).

The change in the capital stock over time is given by equation (1.2). If we divide both sides of this equation by $L$, we get

$$\dot{K}/L = s \cdot f(k) - \delta k$$

The right-hand side contains per capita variables only, but the left-hand side does not. Hence, it is not an ordinary differential equation that can be easily solved. In order to transform it into a differential equation in terms of $k$, we can take the derivative of $k \equiv K/L$ with respect to time to get

$$\dot{k} \equiv \frac{d(K/L)}{dt} = \dot{K}/L - nk$$

where $n = \dot{L}/L$. If we substitute this result into the expression for $\dot{K}/L$, we can rearrange terms to get

$$\dot{k} = s \cdot f(k) - (n+\delta) \cdot k \tag{1.13}$$

Equation (1.13) is the fundamental differential equation of the Solow–Swan model. This nonlinear equation depends only on $k$.

The term $n + \delta$ on the right-hand side of equation (1.13) can be thought of as the effective depreciation rate for the capital-labor ratio, $k \equiv K/L$. If the saving rate, $s$, were 0, capital per person would decline partly due to depreciation of capital at the rate $\delta$ and partly due to the increase in the number of persons at the rate $n$.

Figure 1.1 shows the workings of equation (1.13). The upper curve is the production function, $f(k)$. The term $(n+\delta)\cdot k$, which appears in equation (1.13), is drawn in figure 1.1 as a straight line from the origin with the positive slope $n+\delta$. The term $s\cdot f(k)$ in equation (1.13) looks like the production function except for the multiplication by the positive fraction $s$. Note from the figure that the $s\cdot f(k)$ curve starts from the origin [because $f(0)=0$], has a positive slope [because $f'(k)>0$], and gets flatter as $k$ rises [because $f''(k)<0$]. The Inada conditions imply that the $s\cdot f(k)$ curve is vertical at $k=0$ and becomes flat as $k$ goes to infinity. These properties imply that, other than the origin, the curve $s\cdot f(k)$ and the line $(n+\delta)\cdot k$ cross once and only once.

Consider an economy with the initial capital stock per person $k(0)>0$. Figure 1.1 shows that gross investment per person equals the height of the $s\cdot f(k)$ curve at this point. Consumption per person equals the vertical difference at this point between the $f(k)$ and $s\cdot f(k)$ curves.

### 1.2.3 Markets

In this section we show that the fundamental equation of the Solow–Swan model can be derived in a framework that explicitly incorporates markets. Instead of owning the technology and keeping the output produced with it, we assume that households own financial assets and labor. Assets deliver a rate of return $r(t)$, and labor is paid the wage rate $w(t)$. The total income received by households is, therefore, the sum of asset and labor income, $r(t)\cdot(\text{assets})+w(t)\cdot L(t)$. Households use the income that they do not consume to accumulate more assets

$$d(\text{assets})/dt=[r\cdot(\text{assets})+w\cdot L]-C \tag{1.14}$$

where, again, time subscripts have been omitted to simplify notation. Divide both sides of equation (1.14) by $L$, define assets per person as $a$, and take the derivative of $a$ with respect to time, $\dot{a}=(1/L)\cdot d(\text{assets})/dt-na$, to get that the change in assets per person is given by

$$\dot{a}=(r\cdot a+w)-c-na \tag{1.15}$$

Firms hire labor and capital and use these two inputs with the production technology in equation (1.1) to produce output, which they sell at unit price. We think of firms as renting the services of capital from the households that own it. (None of the results would change if the firms owned the capital, and the households owned shares of stock in the firms.) Hence, the firms' costs of capital are the rental payments, which are proportional to $K$. This specification assumes that capital services can be increased or decreased without incurring any additional expenses, such as costs for installing machines.

Let $R$ be the rental price for a unit of capital services, and assume again that capital stocks depreciate at the constant rate $\delta \geq 0$. The net rate of return to a household that owns a unit of capital is then $R - \delta$. Households also receive the interest rate $r$ on funds lent to other households. In the absence of uncertainty, capital and loans are perfect substitutes as stores of value and, as a result, they must deliver the same return, so $r = R - \delta$ or, equivalently, $R = r + \delta$.

The representative firm's flow of net receipts or profit at any point in time is given by

$$\pi = F(K, L, T) - (r + \delta) \cdot K - wL \tag{1.16}$$

that is, gross receipts from the sale of output, $F(K, L, T)$, less the factor payments, which are rentals to capital, $(r + \delta) \cdot K$, and wages to workers, $wL$. Technology is assumed to be available for free, so no payment is needed to rent the formula used in the process of production. We assume that the firm seeks to maximize the present value of profits. Because the firm rents capital and labor services and has no adjustment costs, there are no intertemporal elements in the firm's maximization problem.[8] (The problem becomes intertemporal when we introduce adjustment costs for capital in chapter 3.)

Consider a firm of arbitrary scale, say with level of labor input $L$. Because the production function exhibits constant returns to scale, the profit for this firm, which is given by equation (1.16), can be written as

$$\pi = L \cdot [f(k) - (r + \delta) \cdot k - w] \tag{1.17}$$

A competitive firm, which takes $r$ and $w$ as given, maximizes profit for given $L$ by setting

$$f'(k) = r + \delta \tag{1.18}$$

That is, the firm chooses the ratio of capital to labor to equate the marginal product of capital to the rental price.

The resulting level of profit is positive, zero, or negative depending on the value of $w$. If profit is positive, the firm could attain infinite profits by choosing an infinite scale. If profit is negative, the firm would contract its scale to zero. Therefore, in a full market equilibrium, $w$ must be such that profit equals zero; that is, the total of the factor payments, $(r + \delta) \cdot K + wL$, equals the gross receipts in equation (1.17). In this case, the firm is indifferent about its scale.

8. In chapter 2 we show that dynamic firms would maximize the present discounted value of all future profits, which is given if $r$ is constant by $\int_0^\infty L \cdot [f(k) - (r + \delta) \cdot k - w] \cdot e^{-rt} dt$. Because the problem does not involve any dynamic constraint, the firm maximizes static profits at all points in time. In fact, this dynamic problem is nothing but a sequence of static problems.

For profit to be zero, the wage rate has to equal the marginal product of labor corresponding to the value of $k$ that satisfies equation (1.18):

$$[f(k) - k \cdot f'(k)] = w \tag{1.19}$$

It can be readily verified from substitution of equations (1.18) and (1.19) into equation (1.17) that the resulting level of profit equals zero for any value of $L$. Equivalently, if the factor prices equal the respective marginal products, the factor payments just exhaust the total output (a result that corresponds in mathematics to Euler's theorem).[9]

The model does not determine the scale of an individual, competitive firm that operates with a constant-returns-to-scale production function. The model will, however, determine the capital/labor ratio $k$, as well as the aggregate level of production, because the aggregate labor force is determined by equation (1.3).

The next step is to define the equilibrium of the economy. In a closed economy, the only asset in positive net supply is capital, because all the borrowing and lending must cancel within the economy. Hence, equilibrium in the asset market requires $a = k$. If we substitute this equality, as well as $r = f'(k) - \delta$ and $w = f(k) - k \cdot f'(k)$, into equation (1.15), we get

$$\dot{k} = f(k) - c - (n + \delta) \cdot k$$

Finally, if we follow Solow–Swan in making the assumption that households consume a constant fraction of their gross income, $c = (1 - s) \cdot f(k)$, we get

$$\dot{k} = s \cdot f(k) - (n + \delta) \cdot k$$

which is the same fundamental equation of the Solow–Swan model that we got in equation (1.13). Hence, introducing competitive markets into the Solow–Swan model does not change any of the main results.[10]

### 1.2.4 The Steady State

We now have the necessary tools to analyze the behavior of the model over time. We first consider the *long run* or *steady state,* and then we describe the *short run* or *transitional dynamics*. We define a *steady state* as a situation in which the various quantities grow at

9. Euler's theorem says that if a function $F(K, L)$ is homogeneous of degree one in $K$ and $L$, then $F(K, L) = F_K \cdot K + F_L \cdot L$. This result can be proven using the equations $F(K, L) = L \cdot f(k)$, $F_K = f'(k)$, and $F_L = f(k) - k \cdot f'(k)$.

10. Note that, in the previous section and here, we assumed that each person saved a constant fraction of his or her gross income. We could have assumed instead that each person saved a constant fraction of his or her net income, $f(k) - \delta k$, which in the market setup equals $ra + w$. In this case, the fundamental equation of the Solow–Swan model would be $\dot{k} = s \cdot f(k) - (s\delta + n) \cdot k$. Again, the same equation applies to the household-producer and market setups.

constant (perhaps zero) rates.[11] In the Solow–Swan model, the steady state corresponds to $\dot{k} = 0$ in equation (1.13),[12] that is, to an intersection of the $s \cdot f(k)$ curve with the $(n+\delta) \cdot k$ line in figure 1.1.[13] The corresponding value of $k$ is denoted $k^*$. (We focus here on the intersection at $k > 0$ and neglect the one at $k = 0$.) Algebraically, $k^*$ satisfies the condition

$$s \cdot f(k^*) = (n+\delta) \cdot k^* \tag{1.20}$$

Since $k$ is constant in the steady state, $y$ and $c$ are also constant at the values $y^* = f(k^*)$ and $c^* = (1-s) \cdot f(k^*)$, respectively. Hence, in the neoclassical model, the per capita quantities $k$, $y$, and $c$ do not grow in the steady state. The constancy of the per capita magnitudes means that the levels of variables—$K$, $Y$, and $C$—grow in the steady state at the rate of population growth, $n$.

Once-and-for-all changes in the level of the technology will be represented by shifts of the production function, $f(\cdot)$. Shifts in the production function, in the saving rate $s$, in the rate of population growth $n$, and in the depreciation rate $\delta$, all have effects on the per capita *levels* of the various quantities in the steady state. In figure 1.1, for example, a proportional upward shift of the production function or an increase in $s$ shifts the $s \cdot f(k)$ curve upward and leads thereby to an increase in $k^*$. An increase in $n$ or $\delta$ moves the $(n+\delta) \cdot k$ line upward and leads to a decrease in $k^*$.

It is important to note that a one-time change in the level of technology, the saving rate, the rate of population growth, and the depreciation rate do not affect the steady-state growth rates of per capita output, capital, and consumption, which are all still equal to zero. For this reason, the model as presently specified will not provide explanations of the determinants of long-run per capita growth.

### 1.2.5 The Golden Rule of Capital Accumulation and Dynamic Inefficiency

For a given level of $A$ and given values of $n$ and $\delta$, there is a unique steady-state value $k^* > 0$ for each value of the saving rate $s$. Denote this relation by $k^*(s)$, with $dk^*(s)/ds > 0$. The steady-state level of per capita consumption is $c^* = (1-s) \cdot f[k^*(s)]$. We know from

11. Some economists use the expression *balanced growth path* to describe the state in which all variables grow at a constant rate and use *steady state* to describe the particular case when the growth rate is zero.

12. We can show that $k$ must be constant in the steady state. Divide both sides of equation (1.13) by $k$ to get $\dot{k}/k = s \cdot f(k)/k - (n+\delta)$. The left-hand side is constant, by definition, in the steady state. Since $s$, $n$, and $\delta$ are all constants, it follows that $f(k)/k$ must be constant in the steady state. The time derivative of $f(k)/k$ equals $-\{[f(k) - kf'(k)]/k\} \cdot (\dot{k}/k)$. The expression $f(k) - kf'(k)$ equals the marginal product of labor (as shown by equation [1.19]) and is positive. Therefore, as long as $k$ is finite, $\dot{k}/k$ must equal 0 in the steady state.

13. The intersection in the range of positive $k$ exists and is unique because $f(0) = 0$, $n+\delta < \lim_{k\to 0}[s \cdot f'(k)] = \infty$, $n+\delta > \lim_{k\to\infty}[s \cdot f'(k)] = 0$, and $f''(k) < 0$.

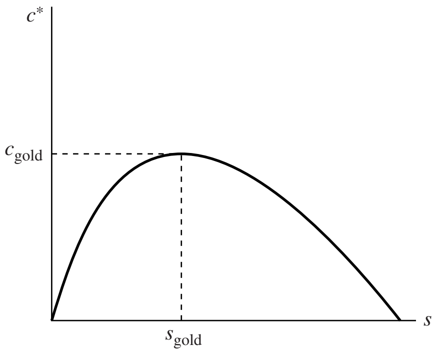


**Figure 1.2**
**The golden rule of capital accumulation.** The vertical axis shows the steady-state level of consumption per person that corresponds to each saving rate. The saving rate that maximizes steady-state consumption per person is called the golden-rule saving rate and is denoted by $s_{\text{Gold}}$.

equation (1.20) that $s \cdot f(k^*) = (n + \delta) \cdot k^*$; hence, we can write an expression for $c^*$ as

$$c^*(s) = f[k^*(s)] - (n + \delta) \cdot k^*(s) \tag{1.21}$$

Figure 1.2 shows the relation between $c^*$ and $s$ that is implied by equation (1.21). The quantity $c^*$ is increasing in $s$ for low levels of $s$ and decreasing in $s$ for high values of $s$. The quantity $c^*$ attains its maximum when the derivative vanishes, that is, when $[f'(k^*) - (n + \delta)] \cdot dk^*/ds = 0$. Since $dk^*/ds > 0$, the term in brackets must equal 0. If we denote the value of $k^*$ that corresponds to the maximum of $c^*$ by $k_{\text{gold}}$, then the condition that determines $k_{\text{gold}}$ is

$$f'(k_{\text{gold}}) = n + \delta \tag{1.22}$$

The corresponding saving rate can be denoted as $s_{\text{gold}}$, and the associated level of steady-state per capita consumption is given by $c_{\text{gold}} = f(k_{\text{gold}}) - (n + \delta) \cdot k_{\text{gold}}$.

The condition in equation (1.22) is called the *golden rule of capital accumulation* (see Phelps, 1966). The source of this name is the biblical Golden Rule, which states, "Do unto others as you would have others do unto you." In economic terms, the golden-rule result can be interpreted as "If we provide the same amount of consumption to members of each current and future generation—that is, if we do not provide less to future generations than to ourselves—then the maximum amount of per capita consumption is $c_{\text{gold}}$."

Figure 1.3 illustrates the workings of the golden rule. The figure considers three possible saving rates, $s_1$, $s_{\text{gold}}$, and $s_2$, where $s_1 < s_{\text{gold}} < s_2$. Consumption per person, $c$, in each case equals the vertical distance between the production function, $f(k)$, and the appropriate

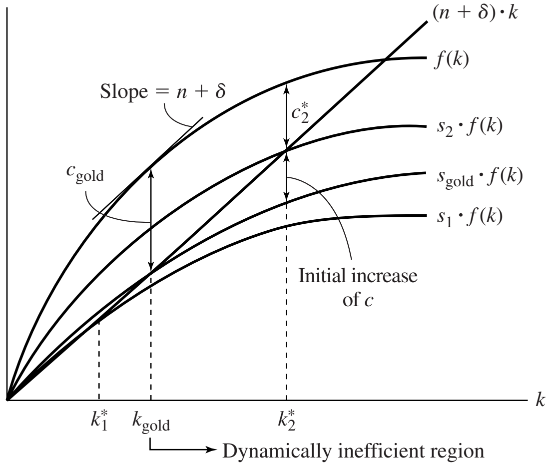


**Figure 1.3**
**The golden rule and dynamic inefficiency.** If the saving rate is above the golden rule ($s_2 > s_{gold}$ in the figure), a reduction in $s$ increases steady-state consumption per person and also raises consumption per person along the transition. Since $c$ increases at all points in time, a saving rate above the golden rule is dynamically inefficient. If the saving rate is below the golden rule ($s_1 < s_{gold}$ in the figure), an increase in $s$ increases steady-state consumption per person but lowers consumption per person along the transition. The desirability of such a change depends on how households trade off current consumption against future consumption.

$s \cdot f(k)$ curve. For each $s$, the steady-state value $k^*$ corresponds to the intersection between the $s \cdot f(k)$ curve and the $(n + \delta) \cdot k$ line. The steady-state per capita consumption, $c^*$, is maximized when $k^* = k_{gold}$ because the tangent to the production function at this point parallels the $(n + \delta) \cdot k$ line. The saving rate that yields $k^* = k_{gold}$ is the one that makes the $s \cdot f(k)$ curve cross the $(n + \delta) \cdot k$ line at the value $k_{gold}$. Since $s_1 < s_{gold} < s_2$, we also see in the figure that $k_1^* < k_{gold} < k_2^*$.

An important question is whether some saving rates are better than others. We will be unable to select the best saving rate (or, indeed, to determine whether a constant saving rate is desirable) until we specify a detailed objective function, as we do in the next chapter. We can, however, argue in the present context that a saving rate that exceeds $s_{gold}$ forever is inefficient because higher quantities of per capita consumption could be obtained at all points in time by reducing the saving rate.

Consider an economy, such as the one described by the saving rate $s_2$ in figure 1.3, for which $s_2 > s_{gold}$, so that $k_2^* > k_{gold}^*$ and $c_2^* < c_{gold}$. Imagine that, starting from the steady state, the saving rate is reduced permanently to $s_{gold}$. Figure 1.3 shows that per capita consumption, $c$—given by the vertical distance between the $f(k)$ and $s_{gold} \cdot f(k)$ curves—initially increases by a discrete amount. Then the level of $c$ falls monotonically during the transition[14] toward its new steady-state value, $c_{gold}$. Since $c_2^* < c_{gold}$, we conclude that $c$ exceeds its previous value, $c_2^*$, at all transitional dates, as well as in the new steady state. Hence, when $s > s_{gold}$, the economy is oversaving in the sense that per capita consumption at all points in time could be raised by lowering the saving rate. An economy that oversaves is said to be *dynamically inefficient,* because the path of per capita consumption lies below feasible alternative paths at all points in time.

If $s < s_{gold}$—as in the case of the saving rate $s_1$ in figure 1.3—then the steady-state amount of per capita consumption can be increased by raising the saving rate. This rise in the saving rate would, however, reduce $c$ currently and during part of the transition period. The outcome will therefore be viewed as good or bad depending on how households weigh today's consumption against the path of future consumption. We cannot judge the desirability of an increase in the saving rate in this situation until we make specific assumptions about how agents discount the future. We proceed along these lines in the next chapter.

### 1.2.6 Transitional Dynamics

The long-run growth rates in the Solow–Swan model are determined entirely by exogenous elements—in the steady state, the per capita quantities $k$, $y$, and $c$ do not grow and the aggregate variables $K$, $Y$, and $C$ grow at the exogenous rate of population growth $n$. Hence, the main substantive conclusions about the long run are that steady-state growth rates are independent of the saving rate or the level of technology. The model does, however, have more interesting implications about transitional dynamics. This transition shows how an economy's per capita income converges toward its own steady-state value and to the per capita incomes of other economies.

Division of both sides of equation (1.13) by $k$ implies that the growth rate of $k$ is given by

$$\gamma_k \equiv \dot{k}/k = s \cdot f(k)/k - (n + \delta) \tag{1.23}$$

where we have used the notation $\gamma_z$ to represent the growth rate of variable $z$, notation that we will use throughout the book. Note that, at all points in time, the growth rate of the level of a variable equals the per capita growth rate plus the exogenous rate of population growth $n$, for example,

$$\dot{K}/K = \dot{k}/k + n$$

For subsequent purposes, we shall find it convenient to focus on the growth rate of $k$, as given in equation (1.23).

14. In the next subsection we analyze the transitional dynamics of the model.

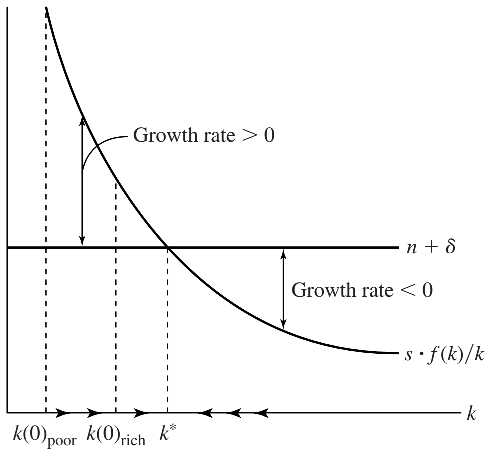


**Figure 1.4**
**Dynamics of the Solow–Swan model.** The growth rate of $k$ is given by the vertical distance between the saving curve, $s \cdot f(k)/k$, and the effective depreciation line, $n + \delta$. If $k < k^*$, the growth rate of $k$ is positive, and $k$ increases toward $k^*$. If $k > k^*$, the growth rate is negative, and $k$ falls toward $k^*$. Thus, the steady-state capital per person, $k^*$, is stable. Note that, along a transition from an initially low capital per person, the growth rate of $k$ declines monotonically toward zero. The arrows on the horizontal axis indicate the direction of movement of $k$ over time.

Equation (1.23) says that $\dot{k}/k$ equals the difference between two terms. The first term, $s \cdot f(k)/k$, we call the *saving curve* and the second term, $(n + \delta)$, the *depreciation curve*. We plot the two curves versus $k$ in figure 1.4. The saving curve is downward sloping;[15] it asymptotes to infinity at $k = 0$ and approaches 0 as $k$ tends to infinity.[16] The depreciation curve is a horizontal line at $n + \delta$. The vertical distance between the saving curve and the depreciation line equals the growth rate of capital per person (from equation [1.23]), and the crossing point corresponds to the steady state. Since $n + \delta > 0$ and $s \cdot f(k)/k$ falls monotonically from infinity to 0, the saving curve and the depreciation line intersect once and only once. Hence, the steady-state capital-labor ratio $k^* > 0$ exists and is unique.

Figure 1.4 shows that, to the left of the steady state, the $s \cdot f(k)/k$ curve lies above $n + \delta$. Hence, the growth rate of $k$ is positive, and $k$ rises over time. As $k$ increases, $\dot{k}/k$ declines and approaches 0 as $k$ approaches $k^*$. (The saving curve gets closer to the depreciation

15. The derivative of $f(k)/k$ with respect to $k$ equals $-[f(k)/k - f'(k)]/k$. The expression in brackets equals the marginal product of labor, which is positive. Hence, the derivative is negative.

16. Note that $\lim_{k\to 0}[s \cdot f(k)/k] = 0/0$. We can apply l'Hôpital's rule to get $\lim_{k\to 0}[s \cdot f(k)/k] = \lim_{k\to 0}[s \cdot f'(k)] = \infty$, from the Inada condition. Similarly, the Inada condition $\lim_{k\to\infty}[f'(k)] = 0$ implies $\lim_{k\to\infty}[s \cdot f(k)/k] = 0$.

line as $k$ gets closer to $k^*$; hence, $\dot{k}/k$ falls.) The economy tends asymptotically toward the steady state in which $k$—and, hence, $y$ and $c$—do not change.

The reason behind the declining growth rates along the transition is the existence of diminishing returns to capital: when $k$ is relatively low, the average product of capital, $f(k)/k$, is relatively high. By assumption, households save and invest a constant fraction, $s$, of this product. Hence, when $k$ is relatively low, the gross investment per unit of capital, $s \cdot f(k)/k$, is relatively high. Capital per worker, $k$, effectively depreciates at the constant rate $n + \delta$. Consequently, the growth rate, $\dot{k}/k$, is also relatively high.

An analogous argument demonstrates that if the economy starts above the steady state, $k(0) > k^*$, then the growth rate of $k$ is negative, and $k$ falls over time. (Note from figure 1.4 that, for $k > k^*$, the $n + \delta$ line lies above the $s \cdot f(k)/k$ curve, and, hence, $\dot{k}/k < 0$.) The growth rate increases and approaches 0 as $k$ approaches $k^*$. Thus, the system is globally stable: for any initial value, $k(0) > 0$, the economy converges to its unique steady state, $k^* > 0$.

We can also study the behavior of output along the transition. The growth rate of output per capita is given by

$$\dot{y}/y = f'(k) \cdot \dot{k}/f(k) = [k \cdot f'(k)/f(k)] \cdot (\dot{k}/k) \tag{1.24}$$

The expression in brackets on the far right is the *capital share,* that is, the share of the rental income on capital in total income.[17]

Equation (1.24) shows that the relation between $\dot{y}/y$ and $\dot{k}/k$ depends on the behavior of the capital share. In the Cobb–Douglas case (equation [1.11]), the capital share is the constant $\alpha$, and $\dot{y}/y$ is the fraction $\alpha$ of $\dot{k}/k$. Hence, the behavior of $\dot{y}/y$ mimics that of $\dot{k}/k$.

More generally, we can substitute for $\dot{k}/k$ from equation (1.23) into equation (1.24) to get

$$\dot{y}/y = s \cdot f'(k) - (n + \delta) \cdot \text{Sh}(k) \tag{1.25}$$

where $\text{Sh}(k) \equiv k \cdot f'(k)/f(k)$ is the capital share. If we differentiate with respect to $k$ and combine terms, we get

$$\partial(\dot{y}/y)/\partial k = \left[\frac{f''(k) \cdot k}{f(k)}\right] \cdot (\dot{k}/k) - \frac{(n + \delta) f'(k)}{f(k)} \cdot [1 - \text{Sh}(k)]$$

Since $0 < \text{Sh}(k) < 1$, the last term on the right-hand side is negative. If $\dot{k}/k \geq 0$, the first term

17. We showed before that, in a competitive market equilibrium, each unit of capital receives a rental equal to its marginal product, $f'(k)$. Hence, $k \cdot f'(k)$ is the income per person earned by owners of capital, and $k \cdot f'(k)/f(k)$—the term in brackets—is the share of this income in total income per person.

on the right-hand side is nonpositive, and, hence, $\partial(\dot{y}/y)/\partial k < 0$. Thus, $\dot{y}/y$ necessarily falls as $k$ rises (and therefore as $y$ rises) in the region in which $\dot{k}/k \geq 0$, that is, if $k \leq k^*$. If $\dot{k}/k < 0$ ($k > k^*$), the sign of $\partial(\dot{y}/y)/\partial k$ is ambiguous for a general form of the production function, $f(k)$. However, if the economy is close to its steady state, the magnitude of $\dot{k}/k$ will be small, and $\partial(\dot{y}/y)/\partial k < 0$ will surely hold even if $k > k^*$.

In the Solow–Swan model, which assumes a constant saving rate, the level of consumption per person is given by $c = (1 - s) \cdot y$. Hence, the growth rates of consumption and income per capita are identical at all points in time, $\dot{c}/c = \dot{y}/y$. Consumption, therefore, exhibits the same dynamics as output.

### 1.2.7 Behavior of Input Prices During the Transition

We showed before that the Solow–Swan framework is consistent with a competitive market economy in which firms maximize profits and households choose to save a constant fraction of gross income. It is interesting to study the behavior of wages and interest rates along the transition as the capital stock increases toward the steady state. We showed that the interest rate equals the marginal product of capital minus the constant depreciation rate, $r = f'(k) - \delta$. Since the interest rate depends on the marginal product of capital, which depends on the capital stock per person, the interest rate moves during the transition as capital changes. The neoclassical production function exhibits diminishing returns to capital, $f''(k) < 0$, so the marginal product of capital declines as capital grows. It follows that the interest rate declines monotonically toward its steady-state value, given by $r^* = f'(k^*) - \delta$.

We also showed that the competitive wage rate was given by $w = f(k) - k \cdot f'(k)$. Again, the wage rate moves as capital increases. To see the behavior of the wage rate, we can take the derivative of $w$ with respect to $k$ to get

$$\frac{\partial w}{\partial k} = f'(k) - f'(k) - k \cdot f''(k) = -k \cdot f''(k) > 0$$

The wage rate, therefore, increases monotonically as the capital stock grows. In the steady state, the wage rate is given by $w^* = f(k^*) - k^* \cdot f'(k^*)$.

The behavior of wages and interest rates can be seen graphically in figure 1.5. The curve shown in the figure is again the production function, $f(k)$. The income per worker received by individual households is given by

$$y = w + Rk \tag{1.26}$$

where $R = r + \delta$ is the rental price of capital. Once the interest rate and the wage rate are determined, $y$ is a linear function of $k$, with intercept $w$ and slope $R$.

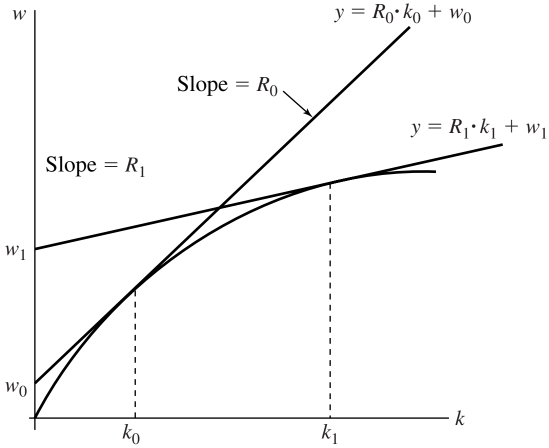


**Figure 1.5**
**Input prices during the transition.** At $k_0$, the straight line that is tangent to the production function has a slope that equals the rental price $R_0$ and an intercept that equals the wage rate $w_0$. As $k$ rises toward $k_1$, the rental price falls toward $R_1$, and the wage rate rises toward $w_1$.

Of course, $R$ depends on $k$ through the marginal productivity condition, $f'(k) = R = r + \delta$. Therefore, $R$, the slope of the income function in equation (1.26), must equal the slope of $f(k)$ at the specified value of $k$. The figure shows two values, $k_0$ and $k_1$. The income functions at these two values are given by straight lines that are tangent to $f(k)$ at $k_0$ and $k_1$, respectively. As $k$ rises during the transition, the figure shows that the slope of the tangent straight line declines from $R_0$ to $R_1$. The figure also shows that the intercept—which equals $w$—rises from $w_0$ to $w_1$.

### 1.2.8 Policy Experiments

Suppose that the economy is initially in a steady-state position with the capital per person equal to $k_1^*$. Imagine that the saving rate rises permanently from $s_1$ to a higher value $s_2$, possibly because households change their behavior or the government introduces some policy that raises the saving rate. Figure 1.6 shows that the $s \cdot f(k)/k$ schedule shifts to the right. Hence, the intersection with the $n + \delta$ line also shifts to the right, and the new steady-state capital stock, $k_2^*$, exceeds $k_1^*$.

How does the economy adjust from $k_1^*$ to $k_2^*$? At $k = k_1^*$, the gap between the $s_1 \cdot f(k)/k$ curve and the $n + \delta$ line is positive; that is, saving is more than enough to generate an increase in $k$. As $k$ increases, its growth rate falls and approaches 0 as $k$ approaches $k_2^*$. The result, therefore, is that a permanent increase in the saving rate generates temporarily

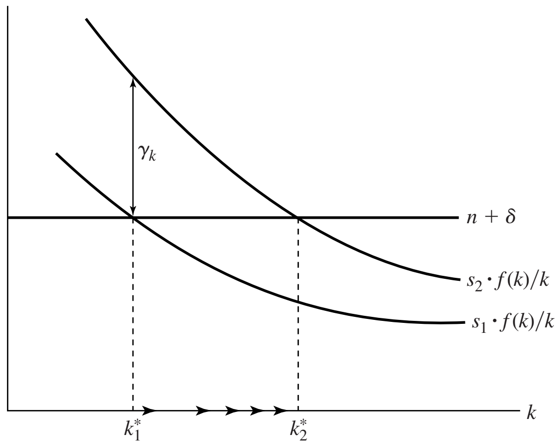


**Figure 1.6**
**Effects from an increase in the saving rate.** Starting from the steady-state capital per person $k_1^*$, an increase in $s$ from $s_1$ to $s_2$ shifts the $s \cdot f(k)/k$ curve to the right. At the old steady state, investment exceeds effective depreciation, and the growth rate of $k$ becomes positive. Capital per person rises until the economy approaches its new steady state at $k_2^* > k_1^*$.

positive per capita growth rates. In the long run, the levels of $k$ and $y$ are permanently higher, but the per capita growth rates return to zero.

The positive transitional growth rates may suggest that the economy could grow forever by raising the saving rate over and over again. One problem with this line of reasoning is that the saving rate is a fraction, a number between zero and one. Since people cannot save more than everything, the saving rate is bounded by one. Notice that, even if people could save all their income, the saving curve would still cross the depreciation line and, as a result, long-run per capita growth would stop.[18] The reason is that the workings of diminishing returns to capital eventually bring the economy back to the zero-growth steady state. Therefore, we can now answer the question that motivated the beginning of this chapter: "Can income per capita grow forever by simply saving and investing physical capital?" If the production function is neoclassical, the answer is "no."

We can also assess permanent changes in the growth rate of population, $n$. These changes could reflect shifts of household behavior or changes in government policies that influence fertility. A decrease in $n$ shifts the depreciation line downward, so that the steady-state level of capital per worker would be larger. However, the long-run growth rate of capital per person would remain at zero.

18. Before reaching $s = 1$, the economy would reach $s_{gold}$, so that further increases in saving rates would put the economy in the dynamically inefficient region.

A permanent, once-and-for-all improvement in the level of the technology has similar, temporary effects on the per capita growth rates. If the production function $f(k)$ shifts upward in a proportional manner, then the saving curve shifts upward, just as in figure 1.6. Hence, $\dot{k}/k$ again becomes positive temporarily. In the long run, the permanent improvement in technology generates higher levels of $k$ and $y$ but no changes in the per capita growth rates. The key difference between improvements in knowledge and increases in the saving rate is that improvements in knowledge are not bounded. That is, the production function can shift over and over again because, in principle, there are no limits to human knowledge. The saving rate, however, is physically bounded by one. It follows that, if we want to generate growth in long-run per capita income and consumption within the neoclassical framework, growth must come from technological progress rather than from physical capital accumulation.

We observed before (note 3) that differences in government policies and institutions can amount to variations in the level of the technology. For example, high tax rates on capital income, failures to protect property rights, and distorting government regulations can be economically equivalent to a poorer level of technology. However, it is probably infeasible to achieve perpetual growth through an unending sequence of improvements in government policies and institutions. Therefore, in the long run, sustained growth would still depend on technological progress.

### 1.2.9 An Example: Cobb–Douglas Technology

We can illustrate the results for the case of a Cobb–Douglas production function (equation [1.11]). The steady-state capital-labor ratio is determined from equation (1.20) as

$$k^* = [sA/(n+\delta)]^{1/(1-\alpha)} \tag{1.27}$$

Note that, as we saw graphically for a more general production function $f(k)$, $k^*$ rises with the saving rate $s$ and the level of technology $A$, and falls with the rate of population growth $n$ and the depreciation rate $\delta$. The steady-state level of output per capita is given by

$$y^* = A^{1/(1-\alpha)} \cdot [s/(n+\delta)]^{\alpha/(1-\alpha)}$$

Thus $y^*$ is a positive function of $s$ and $A$, and a negative function of $n$ and $\delta$.

Along the transition, the growth rate of $k$ is given from equation (1.23) by

$$\dot{k}/k = sAk^{-(1-\alpha)} - (n+\delta) \tag{1.28}$$

If $k(0) < k^*$, then $\dot{k}/k$ in equation (1.28) is positive. This growth rate declines as $k$ rises and approaches 0 as $k$ approaches $k^*$. Since equation (1.24) implies $\dot{y}/y = \alpha \cdot (\dot{k}/k)$, the behavior of $\dot{y}/y$ mimics that of $\dot{k}/k$. In particular, the lower $y(0)$, the higher $\dot{y}/y$.

**A Closed-Form Solution** It is interesting to notice that, when the production function is Cobb–Douglas and the saving rate is constant, it is possible to get a closed-form solution for the exact time path of $k$. Equation (1.28) can be written as

$$\dot{k} \cdot k^{-\alpha} + (n + \delta) \cdot k^{1-\alpha} = sA$$

If we define $v \equiv k^{1-\alpha}$, we can transform the equation to

$$\left(\frac{1}{1-\alpha}\right) \cdot \dot{v} + (n + \delta) \cdot v = sA$$

which is a first-order, linear differential equation in $v$. The solution to this equation is

$$v \equiv k^{1-\alpha} = \frac{sA}{(n+\delta)} + \left\{[k(0)]^{1-\alpha} - \frac{sA}{(n+\delta)}\right\} \cdot e^{-(1-\alpha)\cdot(n+\delta)\cdot t}$$

The last term is an exponential function with exponent equal to $-(1-\alpha) \cdot (n+\delta)$. Hence, the gap between $k^{1-\alpha}$ and its steady-state value, $sA/(n+\delta)$, vanishes exactly at the constant rate $(1-\alpha) \cdot (n+\delta)$.

### 1.2.10 Absolute and Conditional Convergence

The fundamental equation of the Solow–Swan model (equation [1.23]) implies that the derivative of $\dot{k}/k$ with respect to $k$ is negative:

$$\partial(\dot{k}/k)/\partial k = s \cdot [f'(k) - f(k)/k]/k < 0$$

Other things equal, smaller values of $k$ are associated with larger values of $\dot{k}/k$. An important question arises: does this result mean that economies with lower capital per person tend to grow faster in per capita terms? In other words, does there tend to be *convergence* across economies?

To answer these questions, consider a group of closed economies (say, isolated regions or countries) that are structurally similar in the sense that they have the same values of the parameters $s$, $n$, and $\delta$ and also have the same production function $f(\cdot)$. Thus, the economies have the same steady-state values $k^*$ and $y^*$. Imagine that the only difference among the economies is the initial quantity of capital per person $k(0)$. These differences in starting values could reflect past disturbances, such as wars or transitory shocks to production functions. The model then implies that the less-advanced economies—with lower values of $k(0)$ and $y(0)$—have higher growth rates of $k$ and, in the typical case, also higher growth rates of $y$.[19]

19. This conclusion is unambiguous if the production function is Cobb–Douglas, if $k \leq k^*$, or if $k$ is only a small amount above $k^*$.

Figure 1.4 distinguished two economies, one with the low initial value, $k(0)_{\text{poor}}$, and the other with the high initial value, $k(0)_{\text{rich}}$. Since each economy has the same underlying parameters, the dynamics of $k$ are determined in each case by the same $s \cdot f(k)/k$ and $n+\delta$ curves. Hence, the growth rate $\dot{k}/k$ is unambiguously higher for the economy with the lower initial value, $k(0)_{\text{poor}}$. This result implies a form of convergence: regions or countries with lower starting values of the capital-labor ratio have higher per capita growth rates $\dot{k}/k$, and tend thereby to catch up or converge to those with higher capital-labor ratios.

The hypothesis that poor economies tend to grow faster per capita than rich ones—without conditioning on any other characteristics of economies—is referred to as *absolute convergence*. This hypothesis receives only mixed reviews when confronted with data on groups of economies. We can look, for example, at the growth experience of a broad cross section of countries over the period 1960 to 2000. Figure 1.7 plots the average annual growth rate of real per capita GDP against the log of real per capita GDP at the start of the period, 1960, for 114 countries. The growth rates are actually positively correlated with the initial position; that is, there is some tendency for the initially richer countries to grow faster in per capita terms. Thus, this sample rejects the hypothesis of absolute convergence.

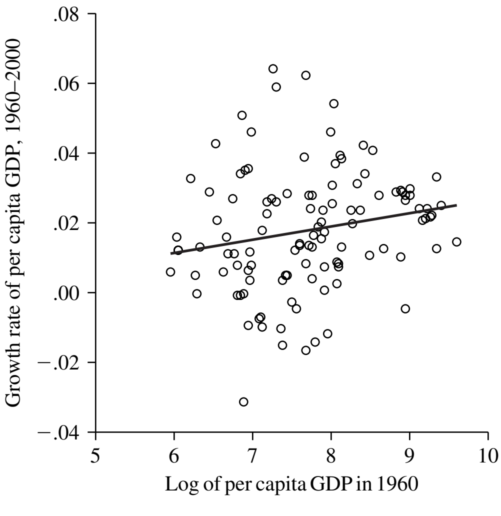


**Figure 1.7**
**Convergence of GDP across countries: Growth rate versus initial level of real per capita GDP for 114 countries.** For a sample of 114 countries, the average growth rate of GDP per capita from 1960 to 2000 (shown on the vertical axis) has little relation with the 1960 level of real per capita GDP (shown on the horizontal axis). The relation is actually slightly positive. Hence, absolute convergence does not apply for a broad cross section of countries.

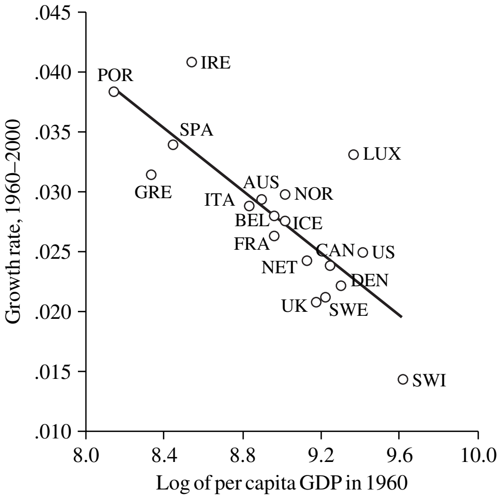


**Figure 1.8**
**Convergence of GDP across OECD countries: Growth rate versus initial level of real per capita GDP for 18 OECD countries.** If the sample is limited to 18 original OECD countries (from 1961), the average growth rate of real per capita GDP from 1960 to 2000 is negatively related to the 1960 level of real per capita GDP. Hence, absolute convergence applies for these OECD countries.

The hypothesis fares better if we examine a more homogeneous group of economies. Figure 1.8 shows the results if we limit consideration to 18 relatively advanced countries that were members of the Organization for Economic Cooperation and Development (OECD) from the start of the organization in 1961.[20] In this case, the initially poorer countries did experience significantly higher per capita growth rates.

This type of result becomes more evident if we consider an even more homogeneous group, the continental U.S. states, each viewed as a separate economy. Figure 1.9 plots the growth rate of per capita personal income for each state from 1880 to 2000 against the log of per capita personal income in 1880.[21] Absolute convergence—the initially poorer states growing faster in per capita terms—holds clearly in this diagram.

We can accommodate the theory to the empirical observations on convergence if we allow for heterogeneity across economies, in particular, if we drop the assumption that all economies have the same parameters, and therefore, the same steady-state positions. If the

20. Germany is omitted because of missing data, and Turkey is omitted because it was not an advanced economy in 1960.

21. There are 47 observations on U.S. states or territories. Oklahoma is omitted because 1880 preceded the Oklahoma land rush, and the data are consequently unavailable.

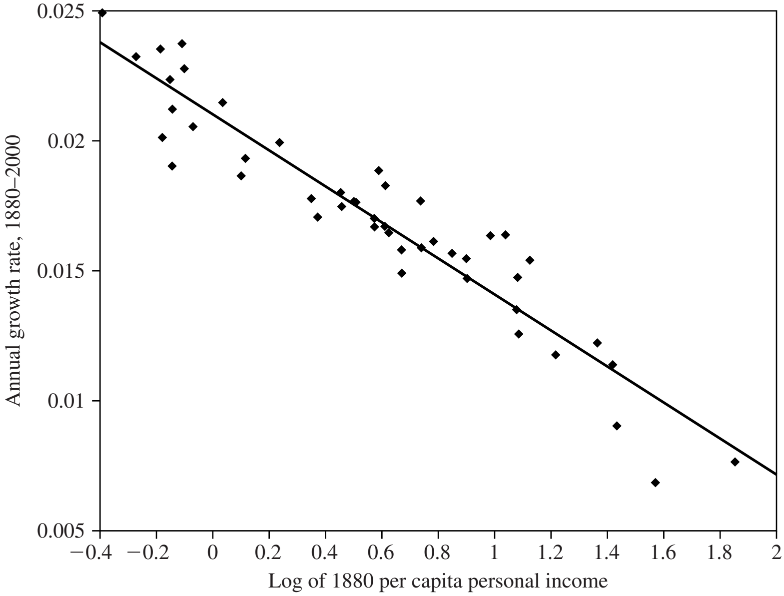


**Figure 1.9**
**Convergence of personal income across U.S. states: 1880 personal income and income growth from 1880 to 2000.** The relation between the growth rate of per capita personal income from 1880 to 2000 (shown on the vertical axis) is negatively related to the level of per capita income in 1880 (shown on the horizontal axis). Thus absolute convergence holds for the states of the United States.

steady states differ, we have to modify the analysis to consider a concept of *conditional convergence*. The main idea is that an economy grows faster the further it is from its own steady-state value.

We illustrate the concept of conditional convergence in figure 1.10 by considering two economies that differ in only two respects: first, they have different initial stocks of capital per person, $k(0)_{\text{poor}} < k(0)_{\text{rich}}$, and second, they have different saving rates, $s_{\text{poor}} \neq s_{\text{rich}}$. Our previous analysis implies that differences in saving rates generate differences in the same direction in the steady-state values of capital per person, that is, $k^*_{\text{poor}} \neq k^*_{\text{rich}}$. [In figure 1.10, these steady-state values are determined by the intersection of the $s_i \cdot f(k)/k$ curves with the common $n + \delta$ line.] We consider the case in which $s_{\text{poor}} < s_{\text{rich}}$ and, hence, $k^*_{\text{poor}} < k^*_{\text{rich}}$ because these differences likely explain why $k(0)_{\text{poor}} < k(0)_{\text{rich}}$ applies at the initial date. (It is also true empirically, as discussed in the introduction, that countries with higher levels of real per capita GDP tend to have higher saving rates.)

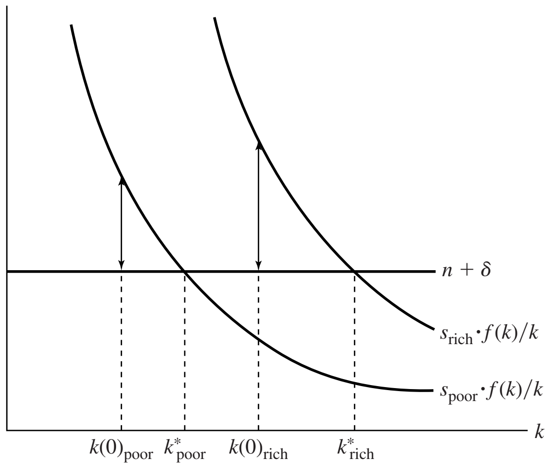


**Figure 1.10**
**Conditional convergence.** If a rich economy has a higher saving rate than a poor economy, the rich economy may be proportionately further from its steady-state position. In this case, the rich economy would be predicted to grow faster per capita than the poor economy; that is, absolute convergence would not hold.

The question is, Does the model predict that the poor economy will grow faster than the rich one? If they have the same saving rate, then the per capita growth rate—the distance between the $s \cdot f(k)/k$ curve and the $n+\delta$ line—would be higher for the poor economy, and $(\dot{k}/k)_{\text{poor}} > (\dot{k}/k)_{\text{rich}}$ would apply. However, if the rich economy has a higher saving rate, as in figure 1.10, then $(\dot{k}/k)_{\text{poor}} < (\dot{k}/k)_{\text{rich}}$ might hold, so that the rich economy grows faster. The intuition is that the low saving rate of the poor economy offsets its higher average product of capital as a determinant of economic growth. Hence, the poor economy may grow at a slower rate than the rich one.

The neoclassical model does predict that each economy converges to its own steady state and that the speed of this convergence relates inversely to the distance from the steady state. In other words, the model predicts conditional convergence in the sense that a lower starting value of real per capita income tends to generate a higher per capita growth rate, once we control for the determinants of the steady state.

Recall that the steady-state value, $k^*$, depends on the saving rate, $s$, and the level of the production function, $f(\cdot)$. We have also mentioned that government policies and institutions can be viewed as additional elements that effectively shift the position of the production function. The findings on conditional convergence suggest that we should hold constant these determinants of $k^*$ to isolate the predicted inverse relationship between growth rates and initial positions.

Algebraically, we can illustrate the concept of conditional convergence by returning to the formula for $\dot{k}/k$ in equation (1.23). One of the determinants of $\dot{k}/k$ is the saving rate $s$. We can use the steady-state condition from equation (1.20) to express $s$ as follows:

$$s = (n + \delta) \cdot k^*/f(k^*)$$

If we replace $s$ by this expression in equation (1.23), then $\dot{k}/k$ can be expressed as

$$\dot{k}/k = (n + \delta) \cdot \left[ \frac{f(k)/k}{f(k^*)/k^*} - 1 \right] \tag{1.29}$$

Equation (1.29) is consistent with $\dot{k}/k = 0$ when $k = k^*$. For given $k^*$, the formula implies that a reduction in $k$, which raises the average product of capital, $f(k)/k$, increases $\dot{k}/k$. But a lower $k$ matches up with a higher $\dot{k}/k$ only if the reduction is relative to the steady-state value, $k^*$. In particular, $f(k)/k$ must be high relative to the steady-state value, $f(k^*)/k^*$. Thus a poor country would not be expected to grow rapidly if its steady-state value, $k^*$, is as low as its current value, $k$.

In the case of a Cobb–Douglas technology, the saving rate can be written as

$$s = \frac{(n + \delta)}{A} \cdot k^{*(1-\alpha)}$$

which we can substitute into equation (1.23) to get

$$\dot{k}/k = (n + \delta) \cdot \left[ \left( \frac{k}{k^*} \right)^{\alpha - 1} - 1 \right] \tag{1.30}$$

We see that the growth rate of capital, $k$, depends on the ratio $k/k^*$; that is, it depends on the distance between the current and steady-state capital-labor ratio.

The result in equation (1.29) suggests that we should look empirically at the relation between the per capita growth rate, $\dot{y}/y$, and the starting position, $y(0)$, after holding fixed variables that account for differences in the steady-state position, $y^*$. For a relatively homogeneous group of economies, such as the U.S. states, the differences in steady-state positions may be minor, and we would still observe the convergence pattern shown in figure 1.9. For a broad cross section of 114 countries, however, as shown in figure 1.7, the differences in steady-state positions are likely to be substantial. Moreover, the countries with low starting levels, $y(0)$, are likely to be in this position precisely because they have low steady-state values, $y^*$, perhaps because of chronically low saving rates or persistently bad government policies that effectively lower the level of the production function. In other words, the per capita growth rate may have little correlation with $\log[y(0)]$, as in figure 1.7, because $\log[y(0)]$ is itself uncorrelated with the gap from the steady state, $\log[y(0)/y^*]$. The perspective of conditional convergence indicates that this gap is the variable that matters for the subsequent per capita growth rate.

We show in chapter 12 that the inclusion of variables that proxy for differences in steady-state positions makes a major difference in the results for the broad cross section of countries. When these additional variables are held constant, the relation between the per capita growth rate and the log of initial real per capita GDP becomes significantly negative, as predicted by the neoclassical model. In other words, the cross-country data support the hypothesis of conditional convergence.

### 1.2.11 Convergence and the Dispersion of Per Capita Income

The concept of convergence considered thus far is that economies with lower levels of per capita income (expressed relative to their steady-state levels of per capita income) tend to grow faster in per capita terms. This behavior is often confused with an alternative meaning of convergence, that the dispersion of real per capita income across a group of economies or individuals tends to fall over time.[22] We show now that, even if absolute convergence holds in our sense, the dispersion of per capita income need not decline over time.

Suppose that absolute convergence holds for a group of economies $i = 1, \ldots, N$, where $N$ is a large number. In discrete time, corresponding for example to annual data, the real per capita income for economy $i$ can then be approximated by the process

$$\log(y_{it}) = a + (1 - b) \cdot \log(y_{i,t-1}) + u_{it} \tag{1.31}$$

where $a$ and $b$ are constants, with $0 < b < 1$, and $u_{it}$ is a disturbance term. The condition $b > 0$ implies absolute convergence because the annual growth rate, $\log(y_{it}/y_{i,t-1})$, is inversely related to $\log(y_{i,t-1})$. A higher coefficient $b$ corresponds to a greater tendency toward convergence.[23] The disturbance term picks up temporary shocks to the production function, the saving rate, and so on. We assume that $u_{it}$ has zero mean, the same variance $\sigma_u^2$ for all economies, and is independent over time and across economies.

One measure of the dispersion or inequality of per capita income is the sample variance of the $\log(y_{it})$:

$$D_t \equiv \frac{1}{N} \cdot \sum_{i=1}^{N} [\log(y_{it}) - \mu_t]^2$$

22. See Sala-i-Martin (1990) and Barro and Sala-i-Martin (1992a) for further discussion of the two concepts of convergence.

23. The condition $b < 1$ rules out a leapfrogging or overshooting effect, whereby an economy that starts out behind another economy would be predicted systematically to get ahead of the other economy at some future date. This leapfrogging effect cannot occur in the neoclassical model but can arise in some models of technological adaptation that we discuss in chapter 8.

where $\mu_t$ is the sample mean of the $\log(y_{it})$. If there are a large number $N$ of observations, the sample variance is close to the population variance, and we can use equation (1.31) to derive the evolution of $D_t$ over time:

$$D_t \approx (1-b)^2 \cdot D_{t-1} + \sigma_u^2$$

This first-order difference equation for dispersion has a steady state given by

$$D^* = \sigma_u^2/[1-(1-b)^2]$$

Hence, the steady-state dispersion falls with $b$ (the strength of the convergence effect) but rises with the variance $\sigma_u^2$ of the disturbance term. In particular, $D^* > 0$ even if $b > 0$, as long as $\sigma_u^2 > 0$.

The evolution of $D_t$ can be expressed as

$$D_t = D^* + (1-b)^2 \cdot (D_{t-1} - D^*) = D^* + (1-b)^{2t} \cdot (D_0 - D^*) \tag{1.32}$$

where $D_0$ is the dispersion at time 0. Since $0 < b < 1$, $D_t$ monotonically approaches its steady-state value, $D^*$, over time. Equation (1.32) implies that $D_t$ rises or falls over time depending on whether $D_0$ begins below or above the steady-state value.[24] Note especially that a rising dispersion is consistent with absolute convergence ($b > 0$).

These results about convergence and dispersion are analogous to Galton's fallacy about the distribution of heights in a population (see Quah, 1993, and Hart, 1995, for discussions). The observation that heights in a family tend to regress toward the mean across generations (a property analogous to our convergence concept for per capita income) does not imply that the dispersion of heights across the full population (a measure that parallels the dispersion of per capita income across economies) tends to narrow over time.

### 1.2.12 Technological Progress

**Classification of Inventions** We have assumed thus far that the level of technology is constant over time. As a result, we found that all per capita variables were constant in the long run. This feature of the model is clearly unrealistic; in the United States, for example, the average per capita growth rate has been positive for over two centuries. In the absence of technological progress, diminishing returns would have made it impossible to maintain per capita growth for so long just by accumulating more capital per worker. The neoclassical economists of the 1950s and 1960s recognized this problem and amended the basic model

24. We could extend the model by allowing for temporary shocks to $\sigma_u^2$ or for major disturbances like wars or oil shocks that affect large subgroups of economies in a common way. In this extended model, the dispersion could depart from the deterministic path that we derived; for example, $D_t$ could rise in some periods even if $D_0$ began above its steady-state value.

to allow the technology to improve over time. These improvements provided an escape from diminishing returns and thus enabled the economy to grow in per capita terms in the long run. We now explore how the model works when we allow for such technological advances.

Although some discoveries are serendipitous, most technological improvements reflect purposeful activity, such as research and development (R&D) carried out in universities and corporate or government laboratories. This research is sometimes financed by private institutions and sometimes by governmental agencies, such as the National Science Foundation. Since the amount of resources devoted to R&D depends on economic conditions, the evolution of the technology also depends on these conditions. This relation will be the subject of our analysis in chapters 6–8. At present, we consider only the simpler case in which the technology improves exogenously.

The first issue is how to introduce exogenous technological progress into the model. This progress can take various forms. Inventions may allow producers to generate the same amount of output with either relatively less capital input or relatively less labor input, cases referred to as *capital-saving* or *labor-saving* technological progress, respectively. Inventions that do not save relatively more of either input are called *neutral* or *unbiased*.

The definition of neutral technological progress depends on the precise meaning of capital saving and labor saving. Three popular definitions are due to Hicks (1932), Harrod (1942), and Solow (1969).

Hicks says that a technological innovation is neutral (Hicks neutral) if the ratio of marginal products remains unchanged for a given capital-labor ratio. This property corresponds to a renumbering of the isoquants, so that Hicks-neutral production functions can be written as

$$Y = T(t) \cdot F(K, L) \tag{1.33}$$

where $T(t)$ is the index of the state of the technology, and $\dot{T}(t) \geq 0$.

Harrod defines an innovation as neutral (Harrod neutral) if the relative input shares, $(K \cdot F_K)/(L \cdot F_L)$, remain unchanged for a given capital-output ratio. Robinson (1938) and Uzawa (1961) showed that this definition implied that the production function took the form

$$Y = F[K, L \cdot T(t)] \tag{1.34}$$

where $T(t)$ is the index of the technology, and $\dot{T}(t) \geq 0$. This form is called *labor-augmenting* technological progress because it raises output in the same way as an increase in the stock of labor. (Notice that the technology factor, $T(t)$, appears in the production function as a multiple of $L$.)

Finally, Solow defines an innovation as neutral (Solow neutral) if the relative input shares, $(L \cdot F_L)/(K \cdot F_K)$, remain unchanged for a given labor/output ratio. This definition can be shown to imply a production function of the form

$$Y = F[K \cdot T(t), L] \tag{1.35}$$

where $T(t)$ is the index of the technology, and $\dot{T}(t) \geq 0$. Production functions of this form are called *capital augmenting* because a technological improvement increases production in the same way as an increase in the stock of capital.

**The Necessity for Technological Progress to Be Labor Augmenting** Suppose that we consider only constant rates of technological progress. Then, in the neoclassical growth model with a constant rate of population growth, only labor-augmenting technological change turns out to be consistent with the existence of a steady state, that is, with constant growth rates of the various quantities in the long run. This result is proved in the appendix to this chapter (section 1.5).

If we want to consider models that possess a steady state, we have to assume that technological progress takes the labor-augmenting form. Another approach, which would be substantially more complicated, would be to deal with models that lack steady states, that is, in which the various growth rates do not approach constants in the long run. However, one reason to stick with the simpler framework that possesses a steady state is that the long-term experiences of the United States and some other developed countries indicate that per capita growth rates can be positive and trendless over long periods of time (see chapter 12). This empirical phenomenon suggests that a useful theory would predict that per capita growth rates approach constants in the long run; that is, the model would possess a steady state.

If the production function is Cobb–Douglas, $Y = AK^{\alpha}L^{1-\alpha}$ in equation (1.11), then it is clear from inspection that the form of technological progress—augmenting $A$, $K$, or $L$—will not matter for the results (see the appendix for discussion). Thus, in the Cobb–Douglas case, we will be safe in assuming that technological progress is labor augmenting. Recall that the key property of the Cobb–Douglas function is that, in a competitive setting, the factor-income shares are constant. Thus, if factor-income shares are reasonably stable—as seems to be true for the U.S. economy but not for some others—we may be okay in regarding the production function as approximately Cobb–Douglas and, hence, in assuming that technogical progress is labor augmenting.

Another approach, when the production function is not Cobb–Douglas, is to derive the form of technological progress from a theory of technological change. Acemoglu (2002) takes this approach, using a variant of the model of endogenous technological change that we develop in chapter 6. He finds that, under some conditions, the form of technological progress would be asymptotically labor augmenting.

**The Solow–Swan Model with Labor-Augmenting Technological Progress** We assume now that the production function includes labor-augmenting technological progress, as shown in equation (1.34), and that the technology term, $T(t)$, grows at the constant rate $x$. The condition for the change in the capital stock is

$$\dot{K} = s \cdot F[K, L \cdot T(t)] - \delta K$$

If we divide both sides of this equation by $L$, we can derive an expression for the change in $k$ over time:

$$\dot{k} = s \cdot F[k, T(t)] - (n + \delta) \cdot k \tag{1.36}$$

The only difference from equation (1.13) is that output per person now depends on the level of the technology, $T(t)$.

Divide both sides of equation (1.36) by $k$ to compute the growth rate:

$$\dot{k}/k = s \cdot F[k, T(t)]/k - (n + \delta) \tag{1.37}$$

As in equation (1.23), $\dot{k}/k$ equals the difference between two terms, where the first term is the product of $s$ and the average product of capital, and the second term is $n + \delta$. The only difference is that now, for given $k$, the average product of capital, $F[k, T(t)]/k$, increases over time because of the growth in $T(t)$ at the rate $x$. In terms of figure 1.4, the downward-sloping curve, $s \cdot F(\cdot)/k$, shifts continually to the right, and, hence, the level of $k$ that corresponds to the intersection between this curve and the $n + \delta$ line also shifts continually to the right. We now compute the growth rate of $k$ in the steady state.

By definition, the steady-state growth rate, $(\dot{k}/k)^*$, is constant. Since $s$, $n$, and $\delta$ are also constants, equation (1.37) implies that the average product of capital, $F[k, T(t)]/k$, is constant in the steady state. Because of constant returns to scale, the expression for the average product equals $F[1, T(t)/k]$ and is therefore constant only if $k$ and $T(t)$ grow at the same rate, that is, $(\dot{k}/k)^* = x$.

Output per capita is given by

$$y = F[k, T(t)] = k \cdot F[1, T(t)/k]$$

Since $k$ and $T(t)$ grow in the steady state at the rate $x$, the steady-state growth rate of $y$ equals $x$. Moreover, since $c = (1 - s) \cdot y$, the steady-state growth rate of $c$ also equals $x$.

To analyze the transitional dynamics of the model with technological progress, it will be convenient to rewrite the system in terms of variables that remain constant in the steady state. Since $k$ and $T(t)$ grow in the steady state at the same rate, we can work with the ratio $\hat{k} \equiv k/T(t) = K/[L \cdot T(t)]$. The variable $L \cdot T(t) \equiv \hat{L}$ is often called the *effective amount of labor*—the physical quantity of labor, $L$, multiplied by its efficiency, $T(t)$. (The terminology *effective labor* is appropriate because the economy operates as if its labor input were $\hat{L}$.) The variable $\hat{k}$ is then the quantity of capital per unit of effective labor.

The quantity of output per unit of effective labor, $\hat{y} \equiv Y/[L \cdot T(t)]$, is given by

$$\hat{y} = F(\hat{k}, 1) \equiv f(\hat{k}) \tag{1.38}$$

Hence, we can again write the production function in intensive form if we replace $y$ and $k$ by $\hat{y}$ and $\hat{k}$, respectively. If we proceed as we did before to get equations (1.13) and (1.23), but now use the condition that $A(t)$ grows at the rate $x$, we can derive the dynamic equation for $\hat{k}$:

$$\dot{\hat{k}}/\hat{k} = s \cdot f(\hat{k})/\hat{k} - (x + n + \delta) \tag{1.39}$$

The only difference between equations (1.39) and (1.23), aside from the hats (ˆ), is that the last term on the right-hand side includes the parameter $x$. The term $x + n + \delta$ is now the effective depreciation rate for $\hat{k} \equiv K/\hat{L}$. If the saving rate, $s$, were zero, $\hat{k}$ would decline partly due to depreciation of $K$ at the rate $\delta$ and partly due to growth of $\hat{L}$ at the rate $x + n$.

Following an argument similar to that of section 1.2.4, we can show that the steady-state growth rate of $\hat{k}$ is zero. The steady-state value $\hat{k}^*$ satisfies the condition

$$s \cdot f(\hat{k}^*) = (x + n + \delta) \cdot \hat{k}^* \tag{1.40}$$

The transitional dynamics of $\hat{k}$ are qualitatively similar to those of $k$ in the previous model. In particular, we can construct a picture like figure 1.4 in which the horizontal axis involves $\hat{k}$, the downward-sloping curve is now $s \cdot f(\hat{k})/\hat{k}$, and the horizontal line is at the level $x + n + \delta$, rather than $n + \delta$. The new construction is shown in figure 1.11. We can use this figure, as we used figure 1.4 before, to assess the relation between the initial value, $\hat{k}(0)$, and the growth rate, $\dot{\hat{k}}/\hat{k}$.

In the steady state, the variables with hats—$\hat{k}$, $\hat{y}$, $\hat{c}$—are now constant. Therefore, the per capita variables—$k, y, c$—now grow in the steady state at the exogenous rate of technological progress, $x$.[25] The level variables—$K$, $Y$, $C$—grow accordingly in the steady state at the rate $n + x$, that is, the sum of population growth and technological change. Note that, as in the prior analysis that neglected technological progress, shifts to the saving rate or the level of the production function affect long-run levels—$\hat{k}^*$, $\hat{y}^*$, $\hat{c}^*$—but not steady-state growth rates. As before, these kinds of disturbances influence growth rates during the transition from an initial position, represented by $\hat{k}(0)$, to the steady-state value, $\hat{k}^*$.

25. We always have the condition $(1/\hat{k}) \cdot (d\hat{k}/dt) = \dot{k}/k - x$. Therefore, $(1/\hat{k}) \cdot (d\hat{k}/dt) = 0$ implies $\dot{k}/k = x$, and similarly for $\dot{y}/y$ and $\dot{c}/c$.

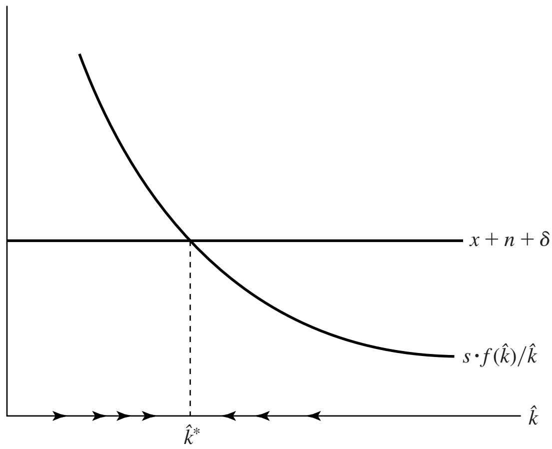


**Figure 1.11**
**The Solow–Swan model with technological progress.** The growth rate of capital per effective worker ($\hat{k} \equiv K/LT$) is given by the vertical distance between the $s \cdot f(\hat{k})/\hat{k}$ curve and the effective depreciation line, $x + n + \delta$. The economy is at a steady state when $\hat{k}$ is constant. Since $T$ grows at the constant rate $x$, the steady-state growth rate of capital per person, $k$, also equals $x$.

### 1.2.13 A Quantitative Measure of the Speed of Convergence

It is important to know the speed of the transitional dynamics. If convergence is rapid, we can focus on steady-state behavior, because most economies would typically be close to their steady states. Conversely, if convergence is slow, economies would typically be far from their steady states, and, hence, their growth experiences would be dominated by the transitional dynamics.

We now provide a quantitative assessment of how fast the economy approaches its steady state for the case of a Cobb–Douglas production function, shown in equation (1.11). (We generalize later to a broader class of production functions.) We can use equation (1.39), with $L$ replaced by $\hat{L}$, to determine the growth rate of $\hat{k}$ in the Cobb–Douglas case as

$$\dot{\hat{k}}/\hat{k} = sA \cdot (\hat{k})^{-(1-\alpha)} - (x + n + \delta) \tag{1.41}$$

The *speed of convergence, $\beta$*, is measured by how much the growth rate declines as the capital stock increases in a proportional sense, that is,

$$\beta \equiv -\frac{\partial(\dot{\hat{k}}/\hat{k})}{\partial \log \hat{k}} \tag{1.42}$$

Notice that we define $\beta$ with a negative sign because the derivative is negative, so that $\beta$ is positive.

To compute $\beta$, we have to rewrite the growth rate in equation (1.41) as a function of $\log(\hat{k})$:

$$\dot{\hat{k}}/\hat{k} = sA \cdot e^{-(1-\alpha)\cdot\log(\hat{k})} - (x + n + \delta) \tag{1.43}$$

We can take the derivative of equation (1.43) with respect to $\log(\hat{k})$ to get an expression for $\beta$:

$$\beta = (1-\alpha)\cdot sA \cdot (\hat{k})^{-(1-\alpha)} \tag{1.44}$$

Notice that the speed of convergence is not constant but, rather, declines monotonically as the capital stock increases toward its steady-state value. At the steady state, $sA \cdot (\hat{k})^{-(1-\alpha)} = (x + n + \delta)$ holds. Therefore, in the neighborhood of the steady state, the speed of convergence equals

$$\beta^* = (1-\alpha)\cdot(x + n + \delta) \tag{1.45}$$

During the transition to the steady state, the convergence rate, $\beta$, exceeds $\beta^*$ but declines over time.

Another way to get the formula for $\beta^*$ is to consider a log-linear approximation of equation (1.41) in the neighborhood of the steady state:

$$\dot{\hat{k}}/\hat{k} \cong -\beta^* \cdot [\log(\hat{k}/\hat{k}^*)] \tag{1.46}$$

where the coefficient $\beta^*$ comes from a log-linearization of equation (1.41) around the steady state. The resulting coefficient can be shown to equal the right-hand side of equation (1.45). See the appendix at the end of this chapter (section 1.5) for the method of derivation of this log-linearization.

Before we consider further the implications of equation (1.45), we will show that it applies also to the growth rate of $\hat{y}$. For a Cobb–Douglas production function, shown in equation (1.11), we have

$$\dot{\hat{y}}/\hat{y} = \alpha \cdot (\dot{\hat{k}}/\hat{k})$$

$$\log(\hat{y}/\hat{y}^*) = \alpha \cdot \log(\hat{k}/\hat{k}^*)$$

If we substitute these formulas into equation (1.46), we get

$$\dot{\hat{y}}/\hat{y} \approx -\beta^* \cdot [\log(\hat{y}/\hat{y}^*)] \tag{1.47}$$

Hence, the convergence coefficient for $\hat{y}$ is the same as that for $\hat{k}$.

The term $\beta^* = (1-\alpha) \cdot (x + n + \delta)$ in equation (1.45) indicates how rapidly an economy's output per effective worker, $\hat{y}$, approaches its steady-state value, $\hat{y}^*$, in the neighborhood of the steady state. For example, if $\beta^* = 0.05$ per year, 5 percent of the gap between $\hat{y}$ and $\hat{y}^*$ vanishes in one year. The half-life of convergence—the time that it takes for half the initial gap to be eliminated—is thus about 14 years.[26] It would take about 28 years for three-quarters of the gap to vanish.

Consider what the theory implies quantitatively about the convergence coefficient, $\beta^* = (1 - \alpha) \cdot (x + n + \delta)$, in equation (1.45). One property is that the saving rate, $s$, does not affect $\beta^*$. This result reflects two offsetting forces that exactly cancel in the Cobb–Douglas case. First, given $\hat{k}$, a higher saving rate leads to greater investment and, therefore, to a faster speed of convergence. Second, a higher saving rate raises the steady-state capital intensity, $\hat{k}^*$, and thereby lowers the average product of capital in the vicinity of the steady state. This effect reduces the speed of convergence. The coefficient $\beta^*$ is also independent of the overall level of efficiency of the economy, $A$. Differences in $A$, like differences in $s$, have two offsetting effects on the convergence speed, and these effects exactly cancel in the Cobb–Douglas case.

To see the quantitative implications of the parameters that enter into equation (1.45), consider the benchmark values $x = 0.02$ per year, $n = 0.01$ per year, and $\delta = 0.05$ per year. These values appear reasonable, for example, for the U.S. economy. The long-term growth rate of real GDP, which is about 2 percent per year, corresponds in the theory to the parameter $x$. The rate of population growth in recent decades is about 1 percent per year, and the measured depreciation rate for the overall stock of structures and equipment is around 5 percent per year.

For given values of the parameters $x$, $n$, and $\delta$, the coefficient $\beta^*$ in equation (1.45) is determined by the capital-share parameter, $\alpha$. A conventional share for the gross income accruing to a narrow concept of physical capital (structures and equipment) is about $\frac{1}{3}$ (see Denison, 1962; Maddison, 1982; and Jorgenson, Gollop, and Fraumeni, 1987). If we use $\alpha = \frac{1}{3}$, equation (1.45) implies $\beta^* = 5.6$ percent per year, which implies a half-life of 12.5 years. In other words, if the capital share is $\frac{1}{3}$, the neoclassical model predicts relatively short transitions.

26. Equation (1.47) is a differential equation in $\log[\hat{y}(t)]$ with the solution

$$\log[\hat{y}(t)] = (1 - e^{-\beta^* t}) \cdot \log(\hat{y}^*) + e^{-\beta^* t} \cdot \log[\hat{y}(0)]$$

The time $t$ for which $\log[\hat{y}(t)]$ is halfway between $\log[\hat{y}(0)]$ and $\log(\hat{y}^*)$ satisfies the condition $e^{-\beta^* t} = 1/2$. The half-life is therefore $\log(2)/\beta^* = 0.69/\beta^*$. Hence, if $\beta^* = 0.05$ per year, the half-life is 14 years.

In chapters 11 and 12 we argue that this predicted speed of convergence is much too high to accord with the empirical evidence. A convergence coefficient, $\beta$, in the range of 1.5 percent to 3.0 percent per year appears to fit better with the data. If $\beta^* = 2.0$ percent per year, the half-life is about 35 years, and the time needed to eliminate three-quarters of an initial gap from the steady-state position is about 70 years. In other words, convergence speeds that are consistent with the empirical evidence imply that the time required for substantial convergence is typically on the order of several generations.

To accord with an observed rate of convergence of about 2 percent per year, the neoclassical model requires a much higher capital-share coefficient. For example, the value $\alpha = 0.75$, together with the benchmark values for the other parameters, implies $\beta^* = 2.0$ percent per year. Although a capital share of 0.75 is too high for a narrow concept of physical capital, this share is reasonable for an expanded measure that also includes human capital.

**An Extended Solow–Swan Model with Physical and Human Capital** One way to increase the capital share is to add human capital to the model. Consider a Cobb–Douglas production function that uses physical capital, $K$, human capital, $H$,[27] and raw labor, $L$:

$$Y = AK^\alpha H^\eta [T(t) \cdot L]^{1-\alpha-\eta} \tag{1.48}$$

where $T(t)$ again grows at the exogenous rate $x$. Divide the production function by $T(t) \cdot L$ to get output per unit of effective labor:

$$\hat{y} = A\hat{k}^\alpha \hat{h}^\eta \tag{1.49}$$

Output can be used on a one-to-one basis for consumption or investment in either type of capital. Following Solow and Swan, we still assume that people consume a constant fraction, $1 - s$, of their gross income, so the accumulation is given by

$$\dot{\hat{k}} + \dot{\hat{h}} = sA\hat{k}^\alpha \hat{h}^\eta - (\delta + n + x) \cdot (\hat{k} + \hat{h}) \tag{1.50}$$

where we have assumed that the two capital goods depreciate at the same constant rate.

The key question is how overall savings will be allocated between physical and human capital. It is reasonable to think that households will invest in the capital good that delivers the higher return, so that the two rates of return—and, hence, the two marginal products of capital—will have to be equated if both forms of investment are taking place. Therefore,

27. Chapters 4 and 5 discuss human capital in more detail.

we have the condition[28]

$$\alpha \cdot \frac{\hat{y}}{\hat{k}} - \delta = \eta \cdot \frac{\hat{y}}{\hat{h}} - \delta \tag{1.51}$$

The equality between marginal products implies a one-to-one relationship between physical and human capital:

$$\hat{h} = \frac{\eta}{\alpha} \cdot \hat{k} \tag{1.52}$$

We can use this relation to eliminate $\hat{h}$ from equation (1.50) to get

$$\dot{\hat{k}} = s\tilde{A}\hat{k}^{\alpha+\eta} - (\delta + n + x) \cdot \hat{k} \tag{1.53}$$

where $\tilde{A} \equiv (\frac{\eta^{\eta}\alpha^{(1-\eta)}}{\alpha+\eta}) \cdot A$ is a constant. Notice that this accumulation equation is the same as equation (1.41), except that the exponent on the capital stock per worker is now the sum of the physical and human capital shares, $\alpha + \eta$, instead of $\alpha$. Using a derivation analogous to that of the previous section, we therefore get an expression for the convergence coefficient in the steady state:

$$\beta^* = (1 - \alpha - \eta) \cdot (\delta + n + x) \tag{1.54}$$

Jorgenson, Gollop, and Fraumeni (1987) estimate a human-capital share of between 0.4 and 0.5. With $\eta = 0.4$ and with the benchmark parameters of the previous section, including $\alpha = \frac{1}{3}$, the predicted speed of convergence would be $\beta^* = 0.021$. Thus, with a broad concept of capital that includes human capital, the Solow–Swan model can generate the rates of convergence that have been observed empirically.

Mankiw, Romer, and Weil (1992) use a production function analogous to equation (1.48). However, instead of making the Solow–Swan assumption that the overall gross saving rate is constant and exogenous, they assume that the investment rates in the two forms of capital are each constant and exogenous. For physical capital, the growth rate is therefore

$$\dot{\hat{k}} = s_k\tilde{A}\hat{k}^{\alpha-1}\hat{h}^{\eta} - (\delta + n + x) = s_k\tilde{A} \cdot e^{-(1-\alpha)\ln\hat{k}} \cdot e^{\eta\ln\hat{h}} - (\delta + n + x) \tag{1.55}$$

28. In a market setup, profit would be $\pi = AK_t^{\alpha}H_t^{\eta}(T_tL_t)^{1-\alpha-\eta} - R_kK - R_hH - wL$, where $R_k$ and $R_h$ are the rental rates of physical and human capital, respectively. The first-order conditions for the firm require that the marginal products of each of the capital goods be equalized to the rental rates, $R_k = \alpha\frac{\hat{y}}{\hat{k}}$ and $R_h = \eta\frac{\hat{y}}{\hat{h}}$. In an environment without uncertainty, like the one we are considering, physical capital, human capital, and loans are perfect substitutes as stores of value and, as a result, their net returns must be the same. In other words, $r = R_k - \delta = R_h - \delta$. Optimizing firms will, therefore, rent physical and human capital up to the point where their marginal products are equal.

where $s_k$ is an exogenous constant. Similarly, for human capital, the growth rate is

$$\dot{\hat{h}} = s_h \tilde{A} \hat{k}^{\alpha} \hat{h}^{\eta - 1} - (\delta + n + x) = s_h \tilde{A} \cdot e^{\alpha \ln \hat{k}} \cdot e^{-(1-\eta) \ln \hat{h}} - (\delta + n + x) \tag{1.56}$$

where $s_h$ is another exogenous constant. A shortcoming of this approach is that the rates of return to physical and human capital are not equated.

The growth rate of $\hat{y}$ is a weighted average of the growth rates of the two inputs:

$$\dot{\hat{y}}/\hat{y} = \alpha \cdot (\dot{\hat{k}}/\hat{k}) + \eta \cdot (\dot{\hat{h}}/\hat{h})$$

If we use equations (1.55) and (1.56) and take a two-dimensional first-order Taylor-series expansion, we get

$$\begin{aligned}\dot{\hat{y}}/\hat{y} = {} & \big[\alpha s_k \tilde{A} \cdot e^{-(1-\alpha)\ln \hat{k}^*} \cdot e^{\eta \ln \hat{h}^*} \cdot [-(1-\alpha)] \\ & + \eta s_h \tilde{A} \cdot e^{\alpha \ln \hat{k}^*} \cdot e^{-(1-\eta)\ln \hat{h}^*} \cdot \alpha\big] \cdot (\ln \hat{k} - \ln \hat{k}^*) \\ & + \big[\alpha s_k \tilde{A} \cdot e^{-(1-\alpha)\ln k^*} \cdot e^{\hat{\eta} \ln h^*} \cdot \eta \\ & + \eta s_h \tilde{A} \cdot e^{\alpha \ln \hat{k}} \cdot e^{-(1-\eta)\ln \hat{h}^*} \cdot [-(1-\eta)]\big] \cdot (\ln \hat{h} - \ln \hat{h}^*)\end{aligned}$$

The steady-state conditions derived from equations (1.55) and (1.56) can be used to get

$$\begin{aligned}\dot{\hat{y}}/\hat{y} &= -(1-\alpha-\eta) \cdot (\delta + n + x) \cdot [\alpha \cdot (\ln \hat{k} - \ln \hat{k}^*) + \eta \cdot (\ln \hat{h} - \ln \hat{h}^*)] \\ &= -\beta^* \cdot (\ln \hat{y} - \ln \hat{y}^*)\end{aligned} \tag{1.57}$$

Therefore, in the neighborhood of the steady state, the convergence coefficient is $\beta^* = (1-\alpha-\eta) \cdot (\delta + n + x)$, just as in equation (1.54).

## 1.3 Models of Endogenous Growth

### 1.3.1 Theoretical Dissatisfaction with Neoclassical Theory

In the mid-1980s it became increasingly clear that the standard neoclassical growth model was theoretically unsatisfactory as a tool to explore the determinants of long-run growth. We have seen that the model without technological change predicts that the economy will eventually converge to a steady state with zero per capita growth. The fundamental reason is the diminishing returns to capital. One way out of this problem was to broaden the concept of capital, notably to include human components, and then assume that diminishing returns did not apply to this broader class of capital. This approach is the one outlined in the next section and explored in detail in chapters 4 and 5. However, another view was that technological progress in the form of the generation of new ideas was the only way that an economy could escape from diminishing returns in the long run. Thus it became a priority to go beyond the treatment of technological progress as exogenous and, instead, to explain this progress within the model of growth. However, endogenous approaches to technological change encountered basic problems within the neoclassical model—the essential reason is the nonrival nature of the ideas that underlie technology.

Remember that a key characteristic of the state of technology, $T$, is that it is a nonrival input to the production process. Hence, the replication argument that we used before to justify the assumption of constant returns to scale suggests that the correct measure of scale is the two rival inputs, capital and labor. Hence, the concept of constant returns to scale that we used is homogeneity of degree one in $K$ and $L$:

$$F(\lambda K, \lambda L, T) = \lambda \cdot F(K, L, T)$$

Recall also that Euler's theorem implies that a function that is homogeneous of degree one can be decomposed as

$$F(K, L, T) = F_K \cdot K + F_L \cdot L \tag{1.58}$$

In our analysis up to this point, we have been assuming that the same technology, $T$, is freely available to all firms. This availability is technically feasible because $T$ is nonrival. However, it may be that $T$ is at least partly excludable—for example, patent protection, secrecy, and experience might allow some producers to have access to technologies that are superior to those available to others. For the moment, we maintain the assumption that technology is nonexcludable, so that all producers have the same access. This assumption also means that a technological advance is immediately available to all producers.

We know from our previous analysis that perfectly competitive firms that take the input prices, $R$ and $w$, as given end up equating the marginal products to the respective input prices, that is, $F_K = R$ and $F_L = w$. It follows from equation (1.58) that the factor payments exhaust the output, so that each firm's profit equals zero at every point in time.

Suppose that a firm has the option to pay a fixed cost, $\kappa$, to improve the technology from $T$ to $T'$. Since the new technology would, by assumption, be freely available to all other producers, we know that the equilibrium values of $R$ and $w$ would again entail a zero flow of profit for each firm. Therefore, the firm that paid the fixed cost, $\kappa$, will end up losing money overall, because the fixed cost would not be recouped by positive profits at any future dates. It follows that the competitive, neoclassical model cannot sustain purposeful investment in technical change if technology is nonexcludable (as well as nonrival).

The obvious next step is to allow the technology to be at least partly excludable. To bring out the problems with this extension, consider the polar case of full excludability, that is, where each firm's technology is completely private. Assume, however, that there are infinitely many ways in which firms can improve knowledge from $T$ to $T'$ by paying the fixed cost $\kappa$—in other words, there is free entry into the business of creating formulas. Suppose that all firms begin with the technology $T$. Would an individual firm then have the incentive to pay $\kappa$ to improve the technology to $T'$? In fact, the incentive appears to be enormous. At the existing input prices, $R$ and $w$, a neoclassical firm with a superior technology would make a pure profit on each unit produced. Because of the assumed constant returns to scale, the firm would be motivated to hire all the capital and labor available in the economy. In this case, the firm would have lots of monopoly power and would likely no longer act as a perfect competitor in the goods and factor markets. So, the assumptions of the competitive model would break down.

A more basic problem with this result is that other firms would have perceived the same profit opportunity and would also have paid the cost $\kappa$ to acquire the better technology, $T'$. However, when many firms improve their technology by the same amount, the competition pushes up the factor prices, $R$ and $w$, so that the flow of profit is again zero. In this case, none of the firms can cover their fixed cost, $\kappa$, just as in the model in which technology was nonexcludable. Therefore, it is not an equilibrium for technological advance to occur (because all innovators make losses) and it is also not an equilibrium for this advance not to occur (because the potential profit to a single innovator is enormous).

These conceptual difficulties motivated researchers to introduce some aspects of imperfect competition to construct satisfactory models in which the level of the technology can be advanced by purposeful activity, such as R&D expenditures. This potential for endogenous technological progress and, hence, *endogenous growth,* may allow an escape from diminishing returns at the aggregate level. Models of this type were pioneered by Romer (1990) and Aghion and Howitt (1992); we consider them in chapters 6–8. For now, we deal only with models in which technology is either fixed or varying in an exogenous manner.

### 1.3.2 The *AK* Model

The key property of this class of endogenous-growth models is the absence of diminishing returns to capital. The simplest version of a production function without diminishing returns is the $AK$ function:[29]

$$Y = AK \tag{1.59}$$

where $A$ is a positive constant that reflects the level of the technology. The global absence of diminishing returns may seem unrealistic, but the idea becomes more plausible if we think of $K$ in a broad sense to include human capital.[30] Output per capita is $y = Ak$, and the average and marginal products of capital are constant at the level $A > 0$.

29. We think that the first economist to use a production function of the $AK$ type was von Neumann (1937).

30. Knight (1944) stressed the idea that diminishing returns might not apply to a broad concept of capital.

If we substitute $f(k)/k = A$ in equation (1.13), we get

$$\dot{k}/k = sA - (n + \delta)$$

We return here to the case of zero technological progress, $x = 0$, because we want to show that per capita growth can now occur in the long run even without exogenous technological change. For a graphical presentation, the main difference is that the downward-sloping saving curve, $s \cdot f(k)/k$, in figure 1.4 is replaced in figure 1.12 by the horizontal line at the level $sA$. The depreciation curve is still the same horizontal line at $n + \delta$. Hence, $\dot{k}/k$ is the vertical distance between the two lines, $sA$ and $n + \delta$. We depict the case in which $sA > (n + \delta)$, so that $\dot{k}/k > 0$. Since the two lines are parallel, $\dot{k}/k$ is constant; in particular, it is independent of $k$. Therefore, $k$ always grows at the steady-state rate, $(\dot{k}/k)^* = sA - (n + \delta)$.

Since $y = Ak$, $\dot{y}/y = \dot{k}/k$ at every point in time. In addition, since $c = (1 - s) \cdot y$, $\dot{c}/c = \dot{k}/k$ also applies. Hence, all the per capita variables in the model always grow at the same, constant rate, given by

$$\gamma^* = sA - (n + \delta) \tag{1.60}$$

Note that an economy described by the $AK$ technology can display positive long-run per capita growth without any technological progress. Moreover, the per capita growth rate

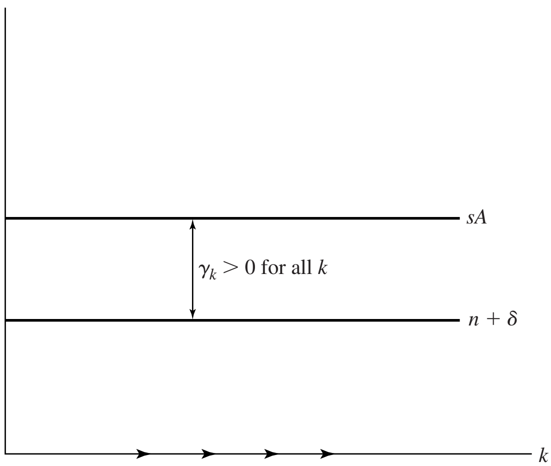


**Figure 1.12**
**The *AK* Model.** If the technology is $AK$, the saving curve, $s \cdot f(k)/k$, is a horizontal line at the level $sA$. If $sA > n + \delta$, perpetual growth of $k$ occurs, even without technological progress.

shown in equation (1.60) depends on the behavioral parameters of the model, including $s$, $A$, and $n$. For example, unlike the neoclassical model, a higher saving rate, $s$, leads to a higher rate of long-run per capita growth, $\gamma^*$.[31] Similarly if the level of the technology, $A$, improves once and for all (or if the elimination of a governmental distortion effectively raises $A$), then the long-run growth rate is higher. Changes in the rates of depreciation, $\delta$, and population growth, $n$, also have permanent effects on the per capita growth rate.

Unlike the neoclassical model, the $AK$ formulation does not predict absolute or conditional convergence, that is, $\partial(\dot{y}/y)/\partial y = 0$ applies for all levels of $y$. Consider a group of economies that are structurally similar in that the parameters $s$, $A$, $n$, and $\delta$ are the same. The economies differ only in terms of their initial capital stocks per person, $k(0)$, and, hence, in $y(0)$ and $c(0)$. Since the model says that each economy grows at the same per capita rate, $\gamma^*$, regardless of its initial position, the prediction is that all the economies grow at the same per capita rate. This conclusion reflects the absence of diminishing returns. Another way to see this result is to observe that the $AK$ model is just a Cobb–Douglas model with a unit capital share, $\alpha = 1$. The analysis of convergence in the previous section showed that the speed of convergence was given in equation (1.45) by $\beta^* = (1 - \alpha) \cdot (x + n + \delta)$; hence, $\alpha = 1$ implies $\beta^* = 0$. This prediction is a substantial failing of the model, because conditional convergence appears to be an empirical regularity. See chapters 11 and 12 for a detailed discussion.

We mentioned that one way to think about the absence of diminishing returns to capital in the $AK$ production function is to consider a broad concept of capital that encompassed physical and human components. In chapters 4 and 5 we consider in more detail models that allow for these two types of capital.

Other approaches have been used to eliminate the tendency for diminishing returns in the neoclassical model. We study in chapter 4 the notion of learning by doing, which was introduced by Arrow (1962) and used by Romer (1986). In these models, the experience with production or investment contributes to productivity. Moreover, the learning by one producer may raise the productivity of others through a process of spillovers of knowledge from one producer to another. Therefore, a larger economy-wide capital stock (or a greater cumulation of the aggregate of past production) improves the level of the technology for each producer. Consequently, diminishing returns to capital may not apply in the aggregate, and increasing returns are even possible. In a situation of increasing returns, each producer's average

31. With the $AK$ production function, we can never get the kind of inefficient oversaving that is possible in the neoclassical model. A shift at some point in time to a permanently higher $s$ means a lower level of $c$ at that point but a permanently higher per capita growth rate, $\gamma^*$, and, hence, higher levels of $c$ after some future date. This change cannot be described as inefficient because it may be desirable or undesirable depending on how households discount future levels of consumption.

product of capital, $f(k)/k$, tends to rise with the economy-wide value of $k$. Consequently, the $s \cdot f(k)/k$ curve in figure 1.4 tends to be upward sloping, at least over some range, and the growth rate, $\dot{k}/k$, rises with $k$ in this range. Thus these kinds of models predict at least some intervals of per capita income in which economies tend to diverge. It is unclear, however, whether these divergence intervals are present in the data.

### 1.3.3 Endogenous Growth with Transitional Dynamics

The $AK$ model delivers endogenous growth by avoiding diminishing returns to capital in the long run. This particular production function also implies, however, that the marginal and average products of capital are always constant and, hence, that growth rates do not exhibit the convergence property. It is possible to retain the feature of constant returns to capital in the long run, while restoring the convergence property—an idea brought out by Jones and Manuelli (1990).[32]

Consider again the expression for the growth rate of $k$ from equation (1.13):

$$\dot{k}/k = s \cdot f(k)/k - (n + \delta) \tag{1.61}$$

If a steady state exists, the associated growth rate, $(\dot{k}/k)^*$, is constant by definition. A positive $(\dot{k}/k)^*$ means that $k$ grows without bound. Equation (1.13) implies that it is necessary and sufficient for $(\dot{k}/k)^*$ to be positive to have the average product of capital, $f(k)/k$, remain above $(n + \delta)/s$ as $k$ approaches infinity. In other words, if the average product approaches some limit, then $\lim_{k\to\infty}[f(k)/k] > (n + \delta)/s$ is necessary and sufficient for endogenous, steady-state growth.

If $f(k) \to \infty$ as $k \to \infty$, then an application of l'Hôpital's rule shows that the limits as $k$ approaches infinity of the average product, $f(k)/k$, and the marginal product, $f'(k)$, are the same. (We assume here that $\lim_{k\to\infty}[f'(k)]$ exists.) Hence, the key condition for endogenous, steady-state growth is that $f'(k)$ be bounded sufficiently far above 0:

$$\lim_{k\to\infty}[f(k)/k] = \lim_{k\to\infty}[f'(k)] > (n + \delta)/s > 0$$

This inequality violates one of the standard Inada conditions in the neoclassical model, $\lim_{k\to\infty}[f'(k)] = 0$. Economically, the violation of this condition means that the tendency for diminishing returns to capital tends to disappear. In other words, the production function can exhibit diminishing or increasing returns to $k$ when $k$ is low, but the marginal product of capital must be bounded from below as $k$ becomes large. A simple example, in which the production function converges asymptotically to the $AK$ form, is

$$Y = F(K, L) = AK + BK^{\alpha}L^{1-\alpha} \tag{1.62}$$

32. See Kurz (1968) for a related discussion.

where $A > 0$, $B > 0$, and $0 < \alpha < 1$. Note that this production function is a combination of the $AK$ and Cobb–Douglas functions. It exhibits constant returns to scale and positive and diminishing returns to labor and capital. However, one of the Inada conditions is violated because $\lim_{K\to\infty}(F_K) = A > 0$.

We can write the function in per capita terms as

$$y = f(k) = Ak + Bk^{\alpha}$$

The average product of capital is given by

$$f(k)/k = A + Bk^{-(1-\alpha)}$$

which is decreasing in $k$ but approaches $A$ as $k$ tends to infinity.

The dynamics of this model can be analyzed with the usual expression from equation (1.13):

$$\dot{k}/k = s \cdot \left[A + Bk^{-(1-\alpha)}\right] - (n + \delta) \tag{1.63}$$

Figure 1.13 shows that the saving curve is downward sloping, and the line $n + \delta$ is horizontal. The difference from figure 1.4 is that, as $k$ goes to infinity, the saving curve in figure 1.13 approaches the positive quantity $sA$, rather than 0. If $sA > n + \delta$, as assumed in the figure, the steady-state growth rate, $(\dot{k}/k)^*$, is positive.

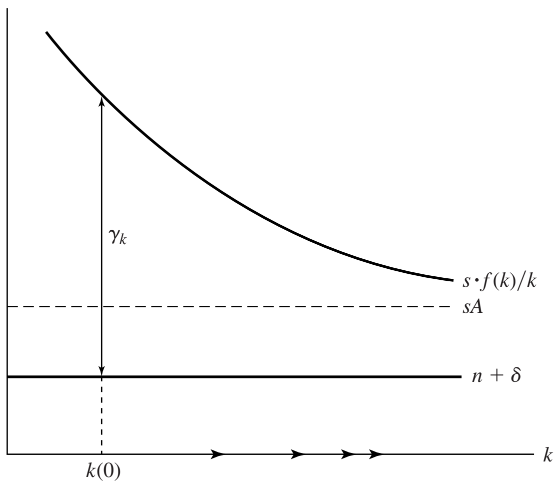


**Figure 1.13**
**Endogenous growth with transitional dynamics.** If the technology is $F(K, L) = AK + BK^{\alpha}L^{1-\alpha}$, the growth rate of $k$ is diminishing for all $k$. If $sA > n+\delta$, the growth rate of $k$ asymptotically approaches a positive constant, given by $sA - n - \delta$. Hence, endogenous growth coexists with a transition in which the growth rate diminishes as the economy develops.

This model yields endogenous, steady-state growth but also predicts conditional convergence, as in the neoclassical model. The reason is that the convergence property derives from the inverse relation between $f(k)/k$ and $k$, a relation that still holds in the model. Figure 1.13 shows that if two economies differ only in terms of their initial values, $k(0)$, the one with the smaller capital stock per person will grow faster in per capita terms.

### 1.3.4 Constant-Elasticity-of-Substitution Production Functions

Consider as another example the production function (due to Arrow et al., 1961) that has a constant elasticity of substitution (CES) between labor and capital:

$$Y = F(K, L) = A \cdot \{a \cdot (bK)^{\psi} + (1-a) \cdot [(1-b) \cdot L]^{\psi}\}^{1/\psi} \tag{1.64}$$

where $0 < a < 1$, $0 < b < 1$,[33] and $\psi < 1$. Note that the production function exhibits constant returns to scale for all values of $\psi$. The elasticity of substitution between capital and labor is $1/(1-\psi)$ (see the appendix, section 1.5.4). As $\psi \to -\infty$, the production function approaches a fixed-proportions technology (discussed in the next section), $Y = \min[bK, (1-b)L]$, where the elasticity of substitution is 0. As $\psi \to 0$, the production function approaches the Cobb–Douglas form, $Y = (\text{constant}) \cdot K^a L^{1-a}$, and the elasticity of substitution is 1 (see the appendix, section 1.5.4). For $\psi = 1$, the production function is linear, $Y = A \cdot [abK + (1-a) \cdot (1-b) \cdot L]$, so that $K$ and $L$ are perfect substitutes (infinite elasticity of substitution).

Divide both sides of equation (1.64) by $L$ to get an expression for output per capita:

$$y = f(k) = A \cdot [a \cdot (bk)^{\psi} + (1-a) \cdot (1-b)^{\psi}]^{1/\psi}$$

The marginal and average products of capital are given, respectively, by

$$f'(k) = Aab^{\psi}[ab^{\psi} + (1-a) \cdot (1-b)^{\psi} \cdot k^{-\psi}]^{(1-\psi)/\psi}$$

$$f(k)/k = A[ab^{\psi} + (1-a) \cdot (1-b)^{\psi} \cdot k^{-\psi}]^{1/\psi}$$

Thus, $f'(k)$ and $f(k)/k$ are each positive and diminishing in $k$ for all values of $\psi$.

We can study the dynamic behavior of a CES economy by returning to the expression from equation (1.13):

$$\dot{k}/k = s \cdot f(k)/k - (n+\delta) \tag{1.65}$$

33. The standard formulation does not include the terms $b$ and $1-b$. The implication then is that the shares of $K$ and $L$ in total product each approach one-half as $\psi \to -\infty$. In our formulation, the shares of $K$ and $L$ approach $b$ and $1-b$, respectively, as $\psi \to -\infty$.

If we graph versus $k$, then $s \cdot f(k)/k$ is a downward-sloping curve, $n + \delta$ is a horizontal line, and $\dot{k}/k$ is still represented by the vertical distance between the curve and the line. The behavior of the growth rate now depends, however, on the parameter $\psi$, which governs the elasticity of substitution between $L$ and $K$.

Consider first the case $0 < \psi < 1$, that is, a high degree of substitution between $L$ and $K$. The limits of the marginal and average products of capital in this case are

$$\lim_{k\to\infty}[f'(k)] = \lim_{k\to\infty}[f(k)/k] = Aba^{1/\psi} > 0$$
$$\lim_{k\to 0}[f'(k)] = \lim_{k\to 0}[f(k)/k] = \infty$$

Hence, the marginal and average products approach a positive constant, rather than 0, as $k$ goes to infinity. In this sense, the CES production function with high substitution between the factors ($0 < \psi < 1$) looks like the example in equation (1.62) in which diminishing returns vanished asymptotically. We therefore anticipate that this CES model can generate endogenous, steady-state growth.

Figure 1.14 shows the results graphically. The $s \cdot f(k)/k$ curve is downward sloping, and it asymptotes to the positive constant $sAb \cdot a^{1/\psi}$. If the saving rate is high enough, so that $sAb \cdot a^{1/\psi} > n+\delta$—as assumed in the figure—then the $s \cdot f(k)/k$ curve always lies above the $n + \delta$ line. In this case, the per capita growth rate is always positive, and the model

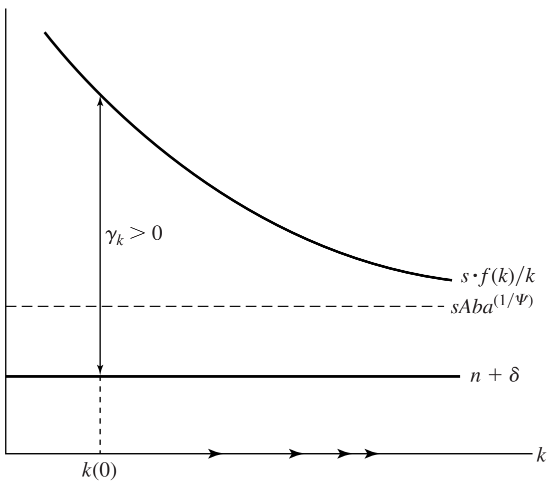


**Figure 1.14**
**The CES model with $0 < \psi < 1$ and $sAb \cdot a^{1/\psi} > n + \delta$.** If the CES technology exhibits a high elasticity of substitution ($0 < \psi < 1$), endogenous growth arises if the parameters satisfy the inequality $sAb \cdot a^{1/\psi} > n + \delta$. Along the transition, the growth rate of $k$ diminishes.

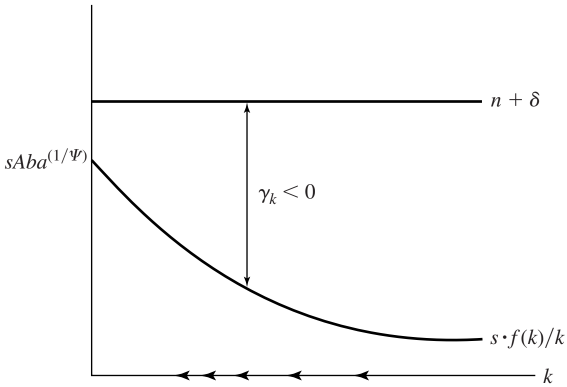


**Figure 1.15**
**The CES model with $\psi < 0$ and $sAb \cdot a^{1/\psi} < n+\delta$.** If the CES technology exhibits a low elasticity of substitution ($\psi < 0$), the growth rate of $k$ would be negative for all levels of $k$ if $sAb \cdot a^{1/\psi} < n+\delta$.

generates endogenous, steady-state growth at the rate

$$\gamma^* = sAb \cdot a^{1/\psi} - (n+\delta)$$

The dynamics of this model are similar to those described in figure 1.13.[34]

Assume now $\psi < 0$, that is, a low degree of substitution between $L$ and $K$. The limits of the marginal and average products of capital in this case are

$$\lim_{k\to\infty}[f'(k)] = \lim_{k\to\infty}[f(k)/k] = 0$$
$$\lim_{k\to 0}[f'(k)] = \lim_{k\to 0}[f(k)/k] = Ab \cdot a^{1/\psi} < \infty$$

Since the marginal and average products approach 0 as $k$ approaches infinity, the key Inada condition is satisfied, and the model does not generate endogenous growth. In this case, however, the violation of the Inada condition as $k$ approaches 0 may cause problems. Suppose that the saving rate is low enough so that $sAb \cdot a^{1/\psi} < n+\delta$. In this case, the $s \cdot f(k)/k$ curve starts at a point below $n+\delta$, and it converges to 0 as $k$ approaches infinity. Figure 1.15 shows, accordingly, that the curve never crosses the $n+\delta$ line, and, hence, no steady state exists with a positive value of $k$. Since the growth rate $\dot{k}/k$ is always negative, the economy shrinks over time, and $k$, $y$, and $c$ all approach 0.[35]

34. If $0 < \psi < 1$ and $sAb \cdot a^{1/\psi} < n+\delta$, then the $s \cdot f(k)/k$ curve crosses $n+\delta$ at the steady-state value $k^*$, as in the standard neoclassical model of figure 1.4. Endogenous growth does not apply in this case.

35. If $\psi < 0$ and $sAb \cdot a^{1/\psi} > n+\delta$, then the $s \cdot f(k)/k$ curve again intersects the $n+\delta$ line at the steady-state value $k^*$.

Since the average product of capital, $f(k)/k$, is a negative function of $k$ for all values of $\psi$, the growth rate $\dot{k}/k$ is also a negative function of $k$. The CES model therefore always exhibits the convergence property: for two economies with identical parameters and different initial values, $k(0)$, the one with the lower value of $k(0)$ has the higher value of $\dot{k}/k$. When the parameters differ across economies, the model predicts conditional convergence, as described before.

We can use the method developed earlier for the case of a Cobb–Douglas production function to derive a formula for the convergence coefficient in the neighborhood of the steady state. The result for a CES production function, which extends equation (1.45), is[36]

$$\beta^* = -(x+n+\delta) \cdot \left[1 - a \cdot \left(\frac{bsA}{x+n+\delta}\right)^{\psi}\right] \tag{1.66}$$

For the Cobb–Douglas case, where $\psi = 0$ and $a = \alpha$, equation (1.66) reduces to equation (1.45). For $\psi \neq 0$, a new result is that $\beta^*$ in equation (1.66) depends on $s$ and $A$. If $\psi > 0$ (high substitutability between $L$ and $K$), then $\beta^*$ falls with $sA$, and vice versa if $\psi < 0$. The coefficient $\beta^*$ is independent of $s$ and $A$ only in the Cobb–Douglas case, where $\psi = 0$.

## 1.4 Other Production Functions . . . Other Growth Theories

### 1.4.1 The Leontief Production Function and the Harrod–Domar Controversy

A production function that was used prior to the neoclassical one is the Leontief (1941), or fixed-proportions, function,

$$Y = F(K, L) = \min(AK, BL) \tag{1.67}$$

where $A > 0$ and $B > 0$ are constants. This specification, which corresponds to $\psi \to -\infty$ in the CES form in equation (1.64), was used by Harrod (1939) and Domar (1946). With fixed proportions, if the available capital stock and labor force happen to be such that $AK = BL$, then all workers and machines are fully employed. If $K$ and $L$ are such that $AK > BL$, then only the quantity of capital $(B/A) \cdot L$ is used, and the remainder remains idle. Conversely, if $AK < BL$, then only the amount of labor $(A/B) \cdot K$ is used, and the remainder is unemployed. The assumption of no substitution between capital and labor led Harrod and Domar to predict that capitalist economies would have undesirable outcomes in the form of perpetual increases in unemployed workers or machines. We provide here a brief analysis of the Harrod–Domar model using the tools developed earlier in this chapter.

36. See Chua (1993) for additional discussion. The formula for $\beta$ in equation (1.66) applies only for cases in which the steady-state level $k^*$ exists. If $0 < \psi < 1$, it applies for $bsA \cdot a^{1/\psi} < x+n+\delta$. If $\psi < 0$, it applies for $bsA \cdot a^{1/\psi} > x+n+\delta$.

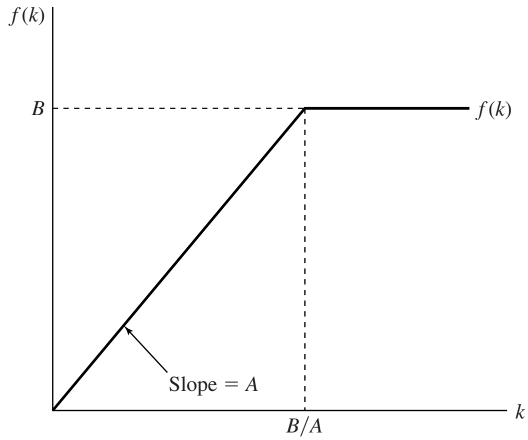


**Figure 1.16**
**The Leontief production function in per capita terms.** In per capita terms, the Leontief production function can be written as $y = \min(Ak, B)$. For $k < B/A$, output per capita is given by $y = Ak$. For $k > B/A$, output per capita is given by $y = B$.

Divide both sides of equation (1.67) by $L$ to get output per capita:

$$y = \min(Ak, B)$$

For $k < B/A$, capital is fully employed, and $y = Ak$. Hence, figure 1.16 shows that the production function in this range is a straight line from the origin with slope $A$. For $k > B/A$, the quantity of capital used is constant, and $Y$ is the constant multiple $B$ of labor, $L$. Hence, output per worker, $y$, equals the constant $B$, as shown by the horizontal part of $f(k)$ in the figure. Note that, as $k$ approaches infinity, the marginal product of capital, $f'(k)$, is zero. Hence, the key Inada condition is satisfied, and we do not expect this production function to yield endogenous steady-state growth.

We can use the expression from equation (1.13) to get

$$\dot{k}/k = s \cdot [\min(Ak, B)]/k - (n + \delta) \tag{1.68}$$

Figures 1.17*a* and 1.17*b* show that the first term, $s \cdot [\min(Ak, B)]/k$, is a horizontal line at $sA$ for $k \leq B/A$. For $k > B/A$, this term is a downward-sloping curve that approaches zero as $k$ goes to infinity. The second term in equation (1.68) is the usual horizontal line at $n + \delta$.

Assume first that the saving rate is low enough so that $sA < n + \delta$, as depicted in figure 1.17. The saving curve, $s \cdot f(k)/k$, then never crosses the $n + \delta$ line, so there is no

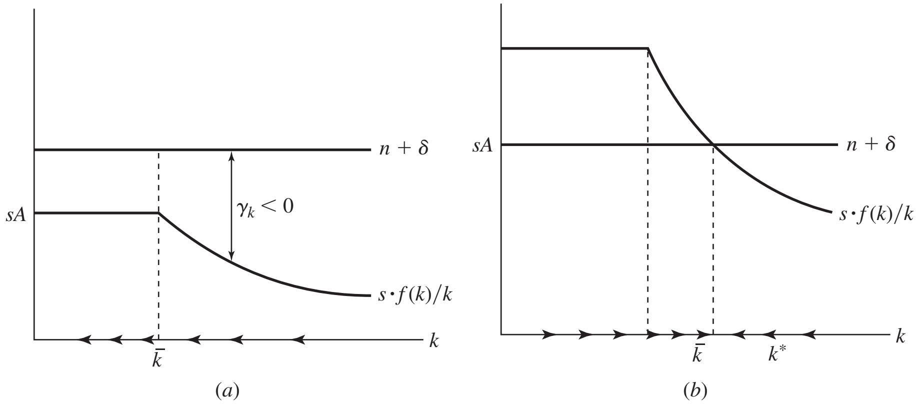


**Figure 1.17**
**The Harrod–Domar model.** In panel *a*, which assumes $sA < n + \delta$, the growth rate of $k$ is negative for all $k$. Therefore, the economy approaches $k = 0$. In panel *b*, which assumes $sA > n + \delta$, the growth rate of $k$ is positive for $k < k^*$ and negative for $k > k^*$, where $k^*$ is the stable steady-state value. Since $k^*$ exceeds $B/A$, a part of the capital stock always remains idle. Moreover, the quantity of idle capital grows steadily (along with $K$ and $L$).

positive steady-state value, $k^*$. Moreover, the growth rate of capital, $\dot{k}/k$, is always negative, so the economy shrinks in per capita terms, and $k$, $y$, and $c$ all approach 0. The economy therefore ends up to the left of $B/A$ and has permanent and increasing unemployment.

Suppose now that the saving rate is high enough so that $sA > n + \delta$, as shown in figure 1.17b. Since the $s \cdot f(k)/k$ curve approaches 0 as $k$ tends to infinity, this curve eventually crosses the $n + \delta$ line at the point $k^* > B/A$. Therefore, if the economy begins at $k(0) < k^*$, $\dot{k}/k$ equals the constant $sA - n - \delta > 0$ until $k$ attains the value $B/A$. At that point, $\dot{k}/k$ falls until it reaches 0 at $k = k^*$. If the economy starts at $k(0) > k^*$, $\dot{k}/k$ is initially negative and approaches 0 as $k$ approaches $k^*$.

Since $k^* > B/A$, the steady state features idle machines but no unemployed workers. Since $k$ is constant in the steady state, the quantity $K$ grows along with $L$ at the rate $n$. Since the fraction of machines that are employed remains constant, the quantity of idle machines also grows at the rate $n$ (yet households are nevertheless assumed to keep saving at the rate $s$).

The only way to reach a steady state in which all capital and labor are employed is for the parameters of the model to satisfy the condition $sA = n + \delta$. Since the four parameters that appear in this condition are all exogenous, there is no reason for the equality to hold. Hence, the conclusion from Harrod and Domar was that an economy would, in all probability, reach one of two undesirable outcomes: perpetual growth of unemployment or perpetual growth of idle machinery.

We know now that there are several implausible assumptions in the arguments of Harrod and Domar. First, the Solow–Swan model showed that Harrod and Domar's parameter $A$—the average product of capital—would typically depend on $k$, and $k$ would adjust to satisfy the equality $s \cdot f(k)/k = n + \delta$ in the steady state. Second, the saving rate could adjust to satisfy this condition. In particular, if agents maximize utility (as we assume in the next chapter), they would not find it optimal to continue to save at the constant rate $s$ when the marginal product of capital was zero. This adjustment of the saving rate would rule out an equilibrium with permanently idle machinery.

### 1.4.2 Growth Models with Poverty Traps

One theme in the literature of economic development concerns *poverty traps*.[37] We can think of a poverty trap as a stable steady state with low levels of per capita output and capital stock. This outcome is a trap because, if agents attempt to break out of it, the economy has a tendency to return to the low-level, stable steady state.

We observed that the average product of capital, $f(k)/k$, declines with $k$ in the neoclassical model. We also noted, however, that this average product may rise with $k$ in some models that feature increasing returns, for example, in formulations that involve learning by doing and spillovers. One way for a poverty trap to arise is for the economy to have an interval of diminishing average product of capital followed by a range of rising average product. (Poverty traps also arise in some models with nonconstant saving rates; see Galor and Ryder, 1989.)

We can get a range of increasing returns by imagining that a country has access to a traditional, as well as a modern, technology.[38] Imagine that producers can use a primitive production function, which takes the usual Cobb–Douglas form,

$$Y_A = AK^{\alpha}L^{1-\alpha} \tag{1.69}$$

The country also has access to a modern, higher productivity technology,[39]

$$Y_B = BK^{\alpha}L^{1-\alpha} \tag{1.70}$$

where $B > A$. However, in order to exploit this better technology, the country as a whole is assumed to have to pay a setup cost at every moment in time, perhaps to cover the necessary public infrastructure or legal system. We assume that this cost is proportional to

37. See especially the *big-push* model of Lewis (1954). A more modern formulation of this idea appears in Murphy, Shleifer, and Vishny (1989).

38. This section is an adaptation of Galor and Zeira (1993), who use two technologies in the context of education.

39. More generally, the capital intensity for the advanced technology would differ from that for the primitive technology. However, this extension complicates the algebra without making any substantive differences.

the labor force and given by $bL$, where $b > 0$. We assume further that this cost is borne by the government and financed by a tax at rate $b$ on each worker. The results are the same whether the tax is paid by producers or workers (who are, in any event, the same persons an economy with household-producers).

In per worker terms, the first production function is

$$y_A = Ak^{\alpha} \tag{1.71}$$

The second production function, when considered net of the setup cost and in per worker terms, is

$$y_B = Bk^{\alpha} - b \tag{1.72}$$

The two production functions are drawn in figure 1.18.

If the government has decided to pay the setup cost, which equals $b$ per worker, all producers will use the modern technology (because the tax $b$ for each worker must be paid in any case). If the government has not paid the setup cost, all producers must use the primitive technology. A sensible government would pay the setup cost if the shift to the modern technology leads to an increase in output per worker at the existing value of $k$ and when measured net of the setup cost. In the present setting, the shift is warranted if $k$ exceeds a critical level, given by $\tilde{k} = [b/(B - A)]^{1/\alpha}$. Thus, the critical value of $k$ rises with the setup cost parameter, $b$, and falls with the difference in the productivity parameters, $B - A$. We assume that the government pays the setup cost if $k \geq \tilde{k}$ and does not pay it if $k < \tilde{k}$.

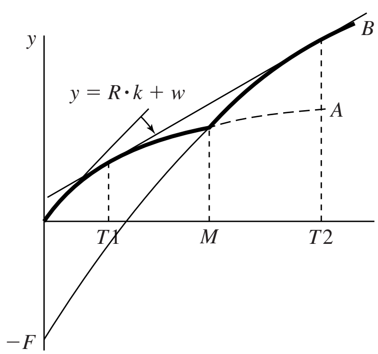


**Figure 1.18**
**Traditional and modern production functions.** The traditional production function has relatively low productivity. The modern production function exhibits higher productivity but is assumed to require a fixed cost to operate.

The growth rate of capital per worker is still given by the fundamental equation of the Solow–Swan model, equation (1.23), as

$$\dot{k}/k = s \cdot f(k)/k - (\delta + n)$$

where $f(k) = Ak^{\alpha}$ if $k < \tilde{k}$ and $f(k) = Bk^{\alpha} - b$ if $k \geq \tilde{k}$. The average product of capital, $f(k)/k$, can be measured graphically in figure 1.18 by the slope of the cord that goes from the origin to the effective production function. We can see that there is a range of $k \geq \tilde{k}$ where the average product is increasing. The saving curve therefore looks like the one depicted in figure 1.19: it has the familiar negative slope at low levels of $k$, is then followed by a range with a positive slope, and again has a negative slope at very high levels of $k$.

Figure 1.19 shows that the $s \cdot f(k)/k$ curve first crosses the $n + \delta$ line at the low steady-state value, $k^*_{low}$, where we assume here that $k^*_{low} < \tilde{k}$. This steady state has the properties that are familiar from the neoclassical model. In particular, $\dot{k}/k > 0$ for $k < k^*_{low}$, and $\dot{k}/k < 0$ at least in an interval of $k > k^*_{low}$. Hence, $k^*_{low}$ is a stable steady state: it is a poverty trap in the sense described before.

The tendency for increasing returns in the middle range of $k$ is assumed to be strong enough so that the $s \cdot f(k)/k$ curve eventually rises to cross the $n + \delta$ line again at the

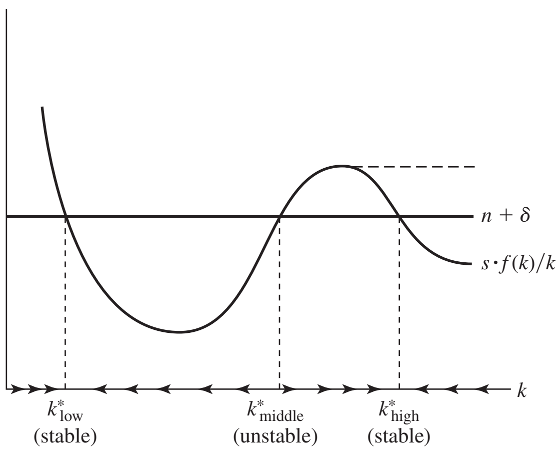


**Figure 1.19**
**A poverty trap.** The production function is assumed to exhibit diminishing returns to $k$ when $k$ is low, increasing returns for a middle range of $k$, and either constant or diminishing returns when $k$ is high. The curve $s \cdot f(k)/k$ is therefore downward sloping for low values of $k$, upward sloping for an intermediate range of $k$, and downward sloping or horizontal for high values of $k$. The steady-state value $k^*_{low}$ is stable and therefore constitutes a poverty trap for countries that begin with $k$ between 0 and $k^*_{middle}$. If a country begins with $k > k^*_{middle}$, it converges to $k^*_{high}$ if diminishing returns to $k$ ultimately set in. If the returns to capital are constant at high values of $k$, as depicted by the dashed portion of the curve, the country converges to a positive long-run growth rate of $k$.

steady-state value $k^*_{\text{middle}}$. This steady state is, however, unstable, because $\dot{k}/k < 0$ applies to the left, and $\dot{k}/k > 0$ holds to the right. Thus, if the economy begins with $k^*_{\text{low}} < k(0) < k^*_{\text{middle}}$, its natural tendency is to return to the development trap at $k^*_{\text{low}}$, whereas if it manages somehow to get to $k(0) > k^*_{\text{middle}}$, it tends to grow further to reach still higher levels of $k$.

In the range where $k > k^*_{\text{middle}}$, the economy's tendency toward diminishing returns eventually brings $s \cdot f(k)/k$ down enough to equal $n + \delta$ at the steady-state value $k^*_{\text{high}}$. This steady state, corresponding to a high level of per capita income but to zero long-term per capita growth, is familiar from our study of the neoclassical model. The key problem for a less-developed economy at the trap level $k^*_{\text{low}}$ is to get over the hump and thereby attain a high long-run level of per capita income.

One empirical implication of the model described by figure 1.19 is that there would exist a middle range of values of $k$—around $k^*_{\text{middle}}$—for which the growth rate, $\dot{k}/k$, is increasing in $k$ and, hence, in $y$. That is, a divergence pattern should hold over this range of per capita incomes. Our reading of the evidence across countries, discussed in chapter 12, does not support this hypothesis. These results are, however, controversial—see, for example, Quah (1996).

## 1.5 Appendix: Proofs of Various Propositions

### 1.5.1 Proof That Each Input Is Essential for Production with a Neoclassical Production Function

We noted in the main body of this chapter that the neoclassical properties of the production function imply that the two inputs, $K$ and $L$, are each essential for production. To verify this proposition, note first that if $Y \to \infty$ as $K \to \infty$, then

$$\lim_{K\to\infty} \frac{Y}{K} = \lim_{K\to\infty} \frac{\partial Y}{\partial K} = 0$$

where the first equality comes from l'Hôpital's rule and the second from the Inada condition. If $Y$ remains bounded as $K$ tends to infinity, then

$$\lim_{K\to\infty} (Y/K) = 0$$

follows immediately. We also know from constant returns to scale that, for any finite $L$,

$$\lim_{K\to\infty} (Y/K) = \lim_{K\to\infty} [F(1, L/K)] = F(1, 0)$$

so that $F(1, 0) = 0$. The condition of constant returns to scale then implies

$$F(K, 0) = K \cdot F(1, 0) = 0$$

for any finite $K$. We can show from an analogous argument that $F(0, L) = 0$ for any finite $L$. These results verify that each input is essential for production.

To demonstrate that output goes to infinity when either input goes to infinity, note that

$$F(K, L) = L \cdot f(k) = K \cdot [f(k)/k]$$

Therefore, for any finite $K$,

$$\lim_{L\to\infty} [F(K, L)] = K \cdot \lim_{k\to 0}[f(k)/k] = K \cdot \lim_{k\to 0}[f'(k)] = \infty$$

where the last equalities follow from l'Hôpital's rule (because essentiality implies $f[0] = 0$) and the Inada condition. We can show from an analogous argument that $\lim_{K\to\infty} [F(K, L)] = \infty$. Therefore, output goes to infinity when either input goes to infinity.

### 1.5.2 Properties of the Convergence Coefficient in the Solow–Swan Model

Equation (1.46) is a log-linearization of equation (1.41) around the steady-state position. To obtain equation (1.46), we have to rewrite equation (1.41) in terms of $\log(\hat{k})$. Note that $\dot{\hat{k}}/\hat{k}$ is the time derivative of $\log(\hat{k})$, and $(\hat{k})^{-(1-\alpha)}$ can be written as $e^{-(1-\alpha)\cdot\log(\hat{k})}$. The steady-state value of $sA(\hat{k})^{-(1-\alpha)}$ equals $x + n + \delta$. We can now take a first-order Taylor expansion of $\log(\hat{k})$ around $\log(\hat{k}^*)$ to get equation (1.46). See the appendix on mathematics at the end of the book for additional discussion. This result appears in Sala-i-Martin (1990) and Mankiw, Romer, and Weil (1992).

The true speed of convergence for $\hat{k}$ or $\hat{y}$ is not constant; it depends on the distance from the steady state. The growth rate of $\hat{y}$ can be written as

$$\dot{\hat{y}}/\hat{y} = \alpha \cdot \left[s \cdot A^{1/\alpha} \cdot (\hat{y})^{-(1-\alpha)/\alpha} - (x + n + \delta)\right]$$

If we use the condition $\hat{y}^* = A \cdot [sA/(x+n+\delta)]^{\alpha/(1-\alpha)}$, we can express the growth rate as

$$\dot{\hat{y}}/\hat{y} = \alpha \cdot (x + n + \delta) \cdot \left[(\hat{y}/\hat{y}^*)^{-(1-\alpha)/\alpha} - 1\right]$$

The convergence coefficient is

$$\beta = -d(\dot{\hat{y}}/\hat{y})]/d[\log(\hat{y})] = (1 - \alpha) \cdot (x + n + \delta) \cdot (\hat{y}/\hat{y}^*)^{-(1-\alpha)/\alpha}$$

At the steady state, $\hat{y} = \hat{y}^*$ and $\beta = (1 - \alpha) \cdot (x + n + \delta)$, as in equation (1.45). More generally, $\beta$ declines as $\hat{y}/\hat{y}^*$ rises.

### 1.5.3 Proof That Technological Progress Must Be Labor Augmenting

We mentioned in the text that technological progress must take the labor-augmenting form shown in equation (1.34) in order for the model to have a steady state with constant growth rates. To prove this result, we start by assuming a production function that includes labor-augmenting and capital-augmenting technological progress:

$$Y = F[K \cdot B(t), L \cdot A(t)] \tag{1.73}$$

where $B(t) = A(t)$ implies that the technological progress is Hicks neutral.

We assume that $A(t) = e^{xt}$ and $B(t) = e^{zt}$, where $x \geq 0$ and $z \geq 0$ are constants. If we divide both sides of equation (1.73) by $K$, we can express output per unit of capital as

$$Y/K = e^{zt} \cdot \left\{ F\left[1, \frac{L \cdot A(t)}{K \cdot B(t)}\right] \right\} = e^{zt} \cdot \varphi\left[(L/K) \cdot e^{(x-z) \cdot t}\right]$$

where $\varphi(\cdot) \equiv F[1, \frac{L \cdot A(t)}{K \cdot B(t)}]$. The population, $L$, grows at the constant rate $n$. If $\gamma_K^*$ is the constant growth rate of $K$ in the steady state, the expression for $Y/K$ can be written as

$$Y/K = e^{zt} \cdot \varphi\left[e^{(n+x-z-\gamma_K^*) \cdot t}\right] \tag{1.74}$$

Recall that the growth rate of $K$ is given by

$$\dot{K}/K = s \cdot (Y/K) - \delta$$

In the steady state, $\dot{K}/K$ equals the constant $\gamma_K^*$, and, hence, $Y/K$ must be constant. There are two ways to get the right-hand side of equation (1.74) to be constant. First, $z = 0$ and $\gamma_K^* = n + x$; that is, technological progress is solely labor augmenting, and the steady-state growth rate of capital equals $n + x$. In this case, the production function can be written in the form of equation (1.34).

The second way to get the right-hand side of equation (1.74) to be constant is with $z \neq 0$ and for the term $\varphi[e^{(n+x-z-\gamma_K^*)t}]$ exactly to offset the term $e^{zt}$. For this case to apply, the derivative of $Y/K$ (in the proposed steady state) with respect to time must be identically zero. If we take the derivative of equation (1.74), set it to zero, and rearrange terms, we get

$$\varphi'(\chi) \cdot \chi/\varphi(\chi) = -z/(n + x - z - \gamma_K^*)$$

where $\chi \equiv e^{(n+x-z-\gamma_K^*) \cdot t}$, and the right-hand side is a constant. If we integrate out, we can write the solution as

$$\varphi(\chi) = (\text{constant}) \cdot \chi^{1-\alpha}$$

where $\alpha$ is a constant. This result implies that the production function can be written as

$$Y = (\text{constant}) \cdot (Ke^{zt})^{\alpha} \cdot (Le^{xt})^{1-\alpha} = (\text{constant}) \cdot K^{\alpha} \cdot (Le^{\nu t})^{1-\alpha}$$

where $\nu = [z\alpha + x \cdot (1 - \alpha)]/(1 - \alpha)$. In other words, if the rate of capital-augmenting technological progress, $z$, is nonzero and a steady state exists, the production function must take the Cobb–Douglas form. Moreover, if the production function is Cobb–Douglas, we can always express technological change as purely labor augmenting (at the rate $\nu$). The conclusion, therefore, is that the existence of a steady state implies that technological progress can be written in the labor-augmenting form.

Another approach to technological progress assumes that capital goods produced later—that is, in a more recent *vintage*—are of higher quality for a given cost. If quality improves in accordance with $T(t)$, the equation for capital accumulation in this vintage model is

$$\dot{K} = s \cdot T(t) \cdot F(K, L) - \delta K \tag{1.75}$$

where $K$ is measured in units of constant quality. This equation corresponds to Hicks-neutral technological progress given by $T(t)$ in the production function. The only difference from the standard specification is that output is $Y = F(K, L)$—not $T(t) \cdot F(K, L)$.

If we want to use a model that possesses a steady state, we would still have to assume that $F(K, L)$ was Cobb–Douglas. In that case, the main properties of the vintage model turn out to be indistinguishable from those of the model that we consider in the text in which technological progress is labor augmenting (see Phelps, 1962, and Solow, 1969, for further discussion). One difference in the vintage model is that, although $K$ and $Y$ grow at constant rates in the steady state, the growth rate of $K$ (in units of constant quality) exceeds that of $Y$. Hence, $K/Y$ is predicted to rise steadily in the long run.

### 1.5.4 Properties of the CES Production Function

The elasticity of substitution is a measure of the curvature of the isoquants. The slope of an isoquant is

$$\frac{dL}{dK}_{\text{isoquant}} = -\frac{\partial F(\cdot)/\partial K}{\partial F(\cdot)/\partial L}$$

The elasticity is given by

$$\left[\frac{\partial(\text{Slope})}{\partial(L/K)} \cdot \frac{L/K}{\text{Slope}}\right]^{-1}$$

For the CES production function shown in equation (1.64), the slope of the isoquant is

$$-(L/K)^{1-\psi} \cdot a \cdot b^{\psi}/[(1-a) \cdot (1-b)^{\psi}]$$

and the elasticity is $1/(1-\psi)$, a constant.

To compute the limit of the production function as $\psi$ approaches 0, use equation (1.64) to get $\lim_{\psi \to 0}[\log(Y)] = \log(A) + 0/0$, which involves an indeterminate form. Apply l'Hôpital's rule to get

$$\begin{aligned}
&\lim_{\psi\to 0}[\log(Y)]\\
&\quad = \log(A) + \left[\frac{a(bK)^{\psi}\cdot\log(bK) + (1-a)\cdot[(1-b)\cdot L]^{\psi}\cdot\log[(1-b)\cdot L]}{a\cdot(bK)^{\psi} + (1-a)\cdot[(1-b)\cdot L]^{\psi}}\right]_{\psi=0}\\
&\quad = \log(A) + a\cdot\log(bK) + (1-a)\cdot\log[(1-b)\cdot L]
\end{aligned}$$

It follows that $Y = \tilde{A}K^{a}L^{1-a}$, where $\tilde{A} = Ab^{a}\cdot(1-b)^{1-a}$. That is, the CES production function approaches the Cobb–Douglas form as $\psi$ tends to zero.

## 1.6 Problems

**1.1 Convergence.**

a. Explain the differences among absolute convergence, conditional convergence, and a reduction in the dispersion of real per capita income across groups.

b. Under what circumstances does absolute convergence imply a decline in the dispersion of per capita income?

**1.2 Forms of technological progress.** Assume that the rate of exogenous technological progress is constant.

a. Show that a steady state can coexist with technological progress only if this progress takes a labor-augmenting form. What is the intuition for this result?

b. Assume that the production function is $Y = F[B(T)\cdot K, A(t)\cdot L]$, where $B(t) = e^{zt}$ and $A(T) = e^{xt}$, with $z \geq 0$ and $x \geq 0$. Show that if $z > 0$ and a steady state exists, the production function must take the Cobb–Douglas form.

**1.3 Dependence of the saving rate, population growth rate, and depreciation rate on the capital intensity.** Assume that the production function satisfies the neoclassical properties.

a. Why would the saving rate, $s$, generally depend on $k$? (Provide some intuition; the precise answer will be given in chapter 2.)

b. How does the speed of convergence change if $s(k)$ is an increasing function of $k$? What if $s(k)$ is a decreasing function of $k$?

Consider now an $AK$ technology.

c. Why would the saving rate, $s$, depend on $k$ in this context?

d. How does the growth rate of $k$ change over time depending on whether $s(k)$ is an increasing or decreasing function of $k$?

e. Suppose that the rate of population growth, $n$, depends on $k$. For an $AK$ technology, what would the relation between $n$ and $k$ have to be in order for the model to predict convergence? Can you think of reasons why $n$ would relate to $k$ in this manner? (We analyze the determination of $n$ in chapter 9.)

f. Repeat part e in terms of the depreciation rate, $\delta$. Why might $\delta$ depend on $k$?

**1.4 Effects of a higher saving rate.** Consider this statement: "Devoting a larger share of national output to investment would help to restore rapid productivity growth and rising living standards." Under what conditions is the statement accurate?

**1.5 Factor shares.** For a neoclassical production function, show that each factor of production earns its marginal product. Show that if owners of capital save all their income and workers consume all their income, the economy reaches the golden rule of capital accumulation. Explain the results.

**1.6 Distortions in the Solow–Swan model (based on Easterly, 1993).** Assume that output is produced by the CES production function,

$$Y = \left[\left(a_F K_F^{\eta} + a_I K_I^{\eta}\right)^{\psi/\eta} + a_G K_G^{\psi}\right]^{1/\psi}$$

where $Y$ is output; $K_F$ is formal capital, which is subject to taxation; $K_I$ is informal capital, which evades taxation; $K_G$ is public capital, provided by government and used freely by all producers; $a_F, a_I, a_G > 0$; $\eta < 1$; and $\psi < 1$. Installed formal and informal capital differ in their location and form of ownership and, therefore, in their productivity.

Output can be used on a one-for-one basis for consumption or gross investment in the three types of capital. All three types of capital depreciate at the rate $\delta$. Population is constant, and technological progress is nil.

Formal capital is subject to tax at the rate $\tau$ at the moment of its installation. Thus, the price of formal capital (in units of output) is $1 + \tau$. The price of a unit of informal capital is one. Gross investment in public capital is the fixed fraction $s_G$ of tax revenues. Any unused tax receipts are rebated to households in a lump-sum manner. The sum of investment in the two forms of private capital is the fraction $s$ of income net of taxes and transfers. Existing private capital can be converted on a one-to-one basis in either direction between formal and informal capital.

a. Derive the ratio of informal to formal capital used by profit-maximizing producers.

b. In the steady state, the three forms of capital grow at the same rate. What is the ratio of output to formal capital in the steady state?

c. What is the steady-state growth rate of the economy?

d. Numerical simulations show that, for reasonable parameter values, the graph of the growth rate against the tax rate, $\tau$, initially increases rapidly, then reaches a peak, and finally decreases steadily. Explain this nonmonotonic relation between the growth rate and the tax rate.

**1.7 A linear production function.** Consider the production function $Y = AK + BL$, where $A$ and $B$ are positive constants.

a. Is this production function neoclassical? Which of the neoclassical conditions does it satisfy and which ones does it not?

b. Write output per person as a function of capital per person. What is the marginal product of $k$? What is the average product of $k$?

In what follows, we assume that population grows at the constant rate $n$ and that capital depreciates at the constant rate $\delta$.

c. Write down the fundamental equation of the Solow–Swan model.

d. Under what conditions does this model have a steady state with no growth of per capita capital, and under what conditions does the model display endogenous growth?

e. In the case of endogenous growth, how does the growth rate of the capital stock behave over time (that is, does it increase or decrease)? What about the growth rates of output and consumption per capita?

f. If $s = 0.4$, $A = 1$, $B = 2$, $\delta = 0.08$, and $n = 0.02$, what is the long-run growth rate of this economy? What if $B = 5$? Explain the differences.

**1.8 Forms of technological progress and steady-state growth.** Consider an economy with a CES production function:

$$Y = D(t) \cdot \{[B(t) \cdot K]^{\psi} + [A(t) \cdot L]^{\psi}\}^{1/\psi}$$

where $\psi$ is a constant parameter different from zero. The terms $D(t)$, $B(t)$, and $A(t)$ represent different forms of technological progress. The growth rates of these three terms are constant, and we denote them by $x_D$, $x_B$, and $x_A$, respectively. Assume that population is constant, with $L = 1$, and normalize the initial levels of the three technologies to one, so that $D(0) = B(0) = A(0) = 1$. In this economy, capital accumulates according to the usual equation:

$$\dot{K} = Y - C - \delta K$$

a. Show that, in a steady state (defined as a situation in which all the variables grow at constant, perhaps different, rates), the growth rates of $Y$, $K$, and $C$ are the same.

b. Imagine first that $x_B = x_A = 0$ and that $x_D > 0$. Show that the steady state must have $\gamma_K = 0$ (and, therefore, $\gamma_Y = \gamma_C = 0$). (Hint: Show first that $\gamma_Y = x_D + \frac{[K_0 e^{\gamma_K t}]^{\psi}}{1+[K_0 e^{\gamma_K t}]^{\psi}} \cdot \gamma_K$.)

c. Using the results in parts a and b, what is the only growth rate of $D(t)$ that is consistent with a steady state? What, therefore, is the only possible steady-state growth rate of $Y$?

d. Imagine now that $x_D = x_A = 0$ and that $x_B > 0$. Show that, in the steady state, $\gamma_K = -x_B$ (Hint: Show first that $\gamma_Y = (x_B + \gamma_K) \cdot \frac{[K_t \cdot B_t]^\psi}{1+[K_t \cdot B_t]^\psi}$.)

e. Using the results in parts a and d, show that the only growth rate of $B$ consistent with a steady state is $x_B = 0$.

f. Finally, assume that $x_D = x_B = 0$ and that $x_A > 0$. Show that, in a steady state, the growth rates must satisfy $\gamma_K = \gamma_Y = \gamma_C = x_D$. (Hint: Show first that $\gamma_Y = \frac{K_t^\psi \cdot \gamma_K + A_t^\psi \cdot x_A}{K_t^\psi + A_t^\psi}$.)

g. What would be the steady-state growth rate in part f if population is not constant but, instead, grows at the rate $n > 0$?